Applied Computational Topology for Point Clouds and Sparse Timeseries Data

Thesis by
Melissa Yeung

In Partial Fulfillment of the Requirements for the
degree of
Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2017
Defended June 8, 2016

ii

ACKNOWLEDGEMENTS

The opportunity to devote uninterrupted time to graduate studies is one of great privilege. A graduate education often marks the beginning of a research career, where one endeavors to expand human knowledge and to make the world a better place for the next generation. Thus, first and foremost, I would like to thank the United States taxpayer, for contributing their hard-earned dollars so that I could have the opportunity to learn, to dream, and to discover.

This dissertation would not have been possible without the support of my advisor, Mathieu Desbrun. Thank you for the freedom to explore, the opportunity to find my own path, for letting me take the scenic route. I also owe much of my graduate experience to Dmitriy Morozov, who has been an indispensable guide and mentor in the world of computational topology and life. Thank you for teaching me the fundamentals—both in computational topology and in high performance computing—and for ensuring that I would always have opportunities—opportunities to present my work, to learn computational topology from the luminaries, and to find belonging in one of the most supportive and welcoming mathematical communities.

The journey that led me to a Ph.D. is improbable. I owe many of my successes along this journey to the numerous individuals who have supported and encouraged me along the way: Miklos Abert, David Beckstead, Collin Bleak, Karen Brucks, Danny Calegari, Gunnar Carlsson, Michael Dorff, Benson Farb, Erica Flapan, Deanna Haunsperger, Stephen Kennedy, David E. Keyes, Robert Maddock, J. Peter May, Konstantin Mischaikow, Paul J. Sally, Jr., Wilhelm Schlag, Stephen Skapek, Liz Stanhope, Glenn Stevens, Elizabeth Townsend, and Samuel L. Volchenboum.

Thanks also to my fellow students for their friendship, support, and wisdom. This journey wouldn't have been nearly as much fun without you.

iv

ABSTRACT

The proliferation of sensors and advancement of technology has led to the production and collection of unprecedented amounts of data in recent years. The data are often noisy, non-linear, and high-dimensional, and the effectiveness of traditional tools may be limited. Thus, the technological advances that enable the ubiquitous collection of data from the cosmological scale to the subatomic scale also necessitate the development of complementary tools that address the new nature of the data.

Recently, there has been much interest in and success with developing topologically-motivated techniques for data analysis. These approaches are especially useful when a topological method is sensitive to large- and small-scale features that might not be detected by methods that require a level of geometric detail that is not provided by the data or by methods that may obscure geometric features, such as principal component analysis (PCA), multidimensional scaling (MDS), and cluster analysis.

Our work explores topological data analysis through two frameworks.

In the first part, we provide a tool for detecting material coherence from a set of spatially sparse particle trajectories via the study of a map induced on homology by the braid corresponding to the motion of particles. While the theory of coherent structures has received a great deal of attention and benefited from many advances in recent years, many of these techniques are limited when the data are sparse. We demonstrate through various examples that our work provides a practical and scalable tool for identifying coherent sets from a sparse set of particle trajectories using eigenanalysis.

In the second part, we formalize the local-to-global structure captured by topology in the setting of point clouds. We extend existing tools in topological data analysis and provide a theoretical framework for studying topological features of a point cloud over a range of resolutions, enabling the analysis of topological features using statistical methods. We apply our tools to the analysis of high-dimensional geospatial sensor data and provide a statistic for quantifying climate anomalies.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

PREFACE

Work for this thesis is motivated by the conviction that our understanding of physical phenomena benefits greatly from a rich intersection of theory from geometry, topology, and dynamical systems. Advances in one field often inspire and bring forth paradigm shifts in another. When I began my Ph.D. studies, I was particularly interested in bringing this synchrony to computational methods for science and engineering applications. This thesis formalizes existing tools of computational topology and introduces new computational methods that allow us to detect dynamical structures from a topological lens.

Topology, since its inception, has been studied and developed as an applied tool for science. Henri Poincaré, often credited as the inventor of algebraic topology, introduced an arsenal of topological techniques and concepts through a series of papers between 1892 and 1904 about the qualitative theory of differential equations and the long-term stability of a mechanical system. Poincaré's topological ideas have been hailed as "probably the greatest advance in celestial mechanics since Newton", as his ideas "not only breathed new life into complex analysis and mechanics; they amounted to the creation of a major new field, algebraic topology" [1].

In the 1930s, Jean Leray (with Juliusz Schauder) developed a set of algebraic tools, including the definition and basic properties of the "topological degree" of a map (related to Brouwer's work), to study fluid dynamics [2]. Leray's subsequent publications throughout the rest of the 1930s provided many applications of topological principles to fluid dynamics and PDEs. *

In recent years, there has been a renewed interest in the interaction of geometry, topology, dynamics, and computation. There is a growing realization

---

*Initially, Leray's interest in algebraic topology was tangential to his other mathematical interests. But in 1940, Leray became a prisoner of the Germans during World War II and spent five years in captivity in an officer's camp, Oflag XVIIA, in Austria. "Leray feared that if his competence as a 'mechanic' ('mécanicien', his word) were known to the German authorities in the camp, he might be compelled to work for the German war machine, so he converted his minor interest to his major one, in fact to his essentially unique one, presented himself as a pure mathematician and devoted himself mainly to algebraic topology" [2].

that one can devise more accurate numerical methods at no additional computational cost by respecting (not just approximating) the appropriate geometric, topological, and dynamical structures at the discrete level. Accordingly, researchers have developed discrete theories of geometry and topology, guided by the tenet that although continuous and discrete mathematics study different structures via different tools, many important properties and relationships can be preserved in the discrete setting under an appropriate framework. The field of discrete differential geometry was developed in this spirit by Mathieu Desbrun, Peter Schröder, and their collaborators [3, 4].

My personal interest in applied topology began with a series of conversations prior to my graduate studies with Gunnar Carlsson at Stanford about his work in a field now commonly referred to as topological data analysis. In his work with coauthors, Carlsson [5] addresses the problem of identifying topological features in high-dimensional point clouds. In a particularly persuasive application of their algorithm, Carlsson and collaborators [6] identify a subgroup of breast cancer patients with excellent survival from microarray data. The subgroup is invisible to traditional methods of data analysis and does not fit into the accepted classification of breast cancers but has a distinct statistically significant molecular signature.

Carlsson's success in applying topological methods of studying microarray data inspired my collaboration with Samuel Volchenboum and Stephen Skapek at the University of Chicago, where we applied topological methods to analyze gene expression profiles of rhabdomyosarcoma tumor samples to predict how a patient might respond to current treatment protocols and identify avenues for future research for therapies that are more effective and less toxic. The time I spent with Volchenboum and Skapek marked one of the most personally and intellectually meaningful periods of my very impressionable younger years and led to my current fervent interest in applied topology.

In graduate school, I had the privilege of learning applied and computational topology from Gunnar Carlsson at Stanford; Dmitriy Morozov at Lawrence Berkeley National Laboratory; Henry Adams (University of Utah), Marko Budišić (University of Wisconsin-Madison), Vin de Silva (Pomona College),

Chapter 1

INTRODUCTION

In this thesis, we study and develop tools in applied computational topology for detecting topological structures in data arising from science and engineering applications.

The computation of the topological invariants of a space, like homology groups from algebraic topology, can offer great utility in situations where metrics and coordinates may not be theoretically justified or when measurements may contain a lot of noise. In these instances, applied computational topology can be leveraged to rigorously solve problems in nonlinear dynamics (possibly from experimental data), data analytics, computer vision, and shape reconstruction. Kaczynski, Mischaikow, and Mrozek [7] provide an overview of some applications of homology to problems involving geometric datasets.

Algebraic topology, though, is not only concerned with the study of topological invariants of a space but also with the study of topological invariants of a representation (e.g., continuous map). Reeb graphs are an example of a topological object that summarizes the structure of a constructible function on a topological space (e.g., a Morse function on a compact manifold or a piecewise linear function on a compact polyhedron) through the compilation of local information (the path-components of level sets).

Continuous maps also induce homomorphisms on homology and cohomology, which can provide a computationally cheap tool for rigorously detecting dynamical features such as the existence of connecting orbits, periodic orbits, and chaotic dynamics. Conley index theory illustrates the power of homology and cohomology for reconciling the continuous dynamics of differential equations with the finite dynamics of computers. Mischaikow [8] provides an overview of the theory in his survey.

Additionally, topology is a powerful tool for obtaining global insights by meaningfully storing and integrating local information. The Euler char-

acteristic is a classic example of a topological invariant that describes the shape and structure of a topological space by counting and summing local information (e.g., number of vertices, edges, and faces).

This local-to-global point of view provided by topology is especially useful in modern applications and with respect to current computational paradigms, where data may be distributed over many local agents (or computers) that may not continuously share information. Accordingly, applied computational topology has found useful applications in robot motion planning, sensor networks, and control theory. We refer readers to the exposition of de Silva and Ghrist [9] for an overview.

Our work interprets these topological ideas for the discrete setting, as we offer new computational methods for detecting dynamical structures as well as extend and obtain theoretical results for existing algorithms in topological data analysis. This thesis consists of two parts.

Overview of Chapter 2: Braids and Material Coherence
In the first part, we study a map induced on homology by the motion of a sparse set of particle trajectories as a tool for detecting material coherence, which refers to temporally coherent structures (or persistent localized features) formed by particle trajectories in the phase space [10]. Material coherence is ubiquitous in nature and plays an important role in science and engineering applications. The importance of coherent structures is evident in the study of solids and fluids, granular flows, molecular dynamics [10], atmospheric and environmental science, propulsion [11], biological defects [12], and even in dynamical systems describing electrical circuits and financial markets [10].

Lagrangian coherent structures are a widely studied example of material coherence. Originally coined by Haller and Yuan [13] to describe the most repelling, attracting, and shearing material surfaces that delineate regions with qualitatively different tracer dynamics, Lagrangian coherent structures provide a simplified understanding of overall flow geometry and quantification of material transport [14]. While there is no universally agreed upon definition for Lagrangian coherent structures, many criteria are based on geometric properties of the flow, such as the finite-time Lyapunov exponent

field [15] or the Cauchy-Green deformation tensor [14, 16–18]. However, these methods are limited when the data are spatially sparse.

Our work focuses on the topological aspects of a flow through the analysis of particle trajectories, and we call upon classical results in topology, like the Nielsen-Thurston classification theorem, in order to provide a new method for detecting coherent sets in two-dimensional flows. By analyzing the flow from a topological perspective, we are able to detect material coherence even when the particle trajectories are spatially sparse. We begin by converting a space-time plot of a set of particle trajectories into a topological braid. The topological braid induces a map on homology, called the (unreduced) Burau representation. We show that coherent sets can be detected as path-components of level sets of an eigenvector of the Burau representation corresponding to the motion of the particles, and we illustrate our method on Aref's blinking vortex flow and a modified Duffing oscillator.

Overview of Chapter 3: Topological Data Analysis
In the second part, we formalize the local-to-global structure captured by topology in the setting of point clouds. We extend existing tools in topological data analysis [5] and provide a theoretical framework for studying topological features of a point cloud over a range of scales. Our work bridges the practical algorithm, mapper, given by Singh, Mémoli, and Carlsson [5], for visualizing a point cloud (potentially given in a very high-dimensional ambient space) as a simplicial complex, and the theoretical work of Munch and Wang [19], which formalizes mapper in a continuous setting (studying path-components of a topological space) in order to prove convergence between mapper constructions and Reeb spaces, a higher-dimensional analogue of Reeb graphs. In our work, we show that mapper constructions over point clouds are stable. More precisely, we show that if D_1 and D_2 are finite sets in X, the distance between their mapper constructions given by a Lipschitz continuous filter function $f : X \to \mathbb{R}^d$ is bounded above by the Hausdorff distance between D_1 and D_2, up to a Lipschitz constant.

To do so, we use category theory to store the data obtained from the mapper algorithm as a functor, and we adapt the interleaving distance given by de Silva et al. [20] in the setting of Reeb graphs and Munch and Wang [19] in

the setting of Reeb spaces to give an analogous interleaving distance in the setting of mapper constructions over point clouds (viewed as functors). We show that dendrograms arising from single-linkage hierarchical clustering can be given as a mapper construction, and in this case, the interleaving distance coincides with the Gromov-Hausdorff distance between dendrograms [21]. As a result, stability and convergence results established by Carlsson and Mémoli for the Gromov-Hausdorff distance between dendrograms applies also to the interleaving distance between dendrograms.

Finally, we introduce the hierarchical complex, which facilitates the study of topological features captured by mapper constructions over a range of resolutions. We construct a hierarchical complex for daily sea surface temperature measurements taken over the course of several decades, and we use persistent homology to provide a statistic for quantifying climate anomalies.

Audience

We hope that our work will appeal to an interdisciplinary audience. As such, we include much background material. We encourage those who are familiar with the background material to skip sections when appropriate.

Chapter 2

BRAIDS AND MATERIAL COHERENCE

We agree on a simplifying assumption that the earth has the shape of a torus.[*]

—Vladimir Igorevich Arnold [23]

2.1 Introduction

Arnold [25] launched the field of topological fluid dynamics in 1966, when he showed that the Euler equations of motion of an ideal incompressible fluid on a Riemannian manifold M can be viewed as geodesic flows on the group of measure-preserving diffeomorphisms of the domain M. Arnold's idea of employing groups to study fluid flows shifted the paradigm of theoretical fluid dynamics and inspired much interest and progress in the field.

Recent work has applied methods from the theory of braid groups to the analysis of mixing in flows in a variety of ways. Motivated by Aref and Pomphrey's [26, 27] study of advection by point vortices on the infinite plane, which laid the groundwork for Aref's seminal paper on chaotic advection [28], where the blinking vortex flow was introduced, Boyland et al. [29] applied Nielsen-Thurston theory to study the motions of systems of point vortices in the infinite plane. These point vortices act as stirrers that displace the fluid. Boyland et al. use the Lagrangian motions of the point vortices to study the Lagrangian motions of the surrounding fluid particles. The motion of the vortices as they wrap around one another can be described using

---

[*]In the same 1966 paper that launched the field of topological fluid dynamics, Arnold also published the first computations of curvatures for diffeomorphism groups [22]. Negative curvature implies exponential instability of geodesics [23]. So making the assumptions that "the earth has the shape of a torus obtained by factoring the plane by a square lattice" and modeling the atmosphere as a two-dimensional homogeneous non-compressible non-viscous fluid, Arnold explains that long-term weather predictions are inherently inaccurate. In particular, a weather forecast two months in advance requires initial data with five more digits of accuracy than the prediction accuracy.

Years later, Lukatskii [24] computes curvatures for diffeomorphism groups on the sphere. Remarkably, his calculations give the same order of magnitude for the error of weather forecasting.

Artin's braid group, which offers a precise framework for distinguishing different regimes of vortex dynamics.

In a related work, Boyland et al. [30] study mixing of a viscous fluid by a (periodic) stirring motion of a finite number of physical rods. The stirring motion of n rods trace out a braid on n strands. Nielsen-Thurston theory can then be applied to give a lower bound on topological entropy. Systems with positive topological entropy exhibit chaotic trajectories.

Gouillart et al. [31] extend the work of Boyland et al. and study stirring protocols in which the motion of the stirring rods is topologically trivial but yet gives rise to a flow with positive topological entropy. In this setting, Nielsen-Thurston theory is applied to the braid formed by considering periodic orbits of the flow in conjuction with the orbit of the stirrer. These periodic orbits act as obstacles to material lines and thus can be considered "ghost rods".

The study of homeomorphisms of a surface punctured by periodic orbits is very classical. The ideas date back to Bowden [32], and we discuss some of these ideas in the context of mapping class groups later.

Thiffeault [33] proposes that the braiding exhibited by the motion of non-periodic orbits can also be used to reveal the presence of topological chaos. To characterize the complexity of the motion, he defines the braiding exponent in terms of the logarithm of the spectral radius of the Burau representation of the braid. It was shown by Fried [34] and Kolev [35] that the topological entropy of a braid, which is related to the Lyapunov exponent of the flow, is bounded below by the logarithm of the spectral radius of the braid's Burau matrix. Thiffeault shows experimentally that the magnitude of the braiding exponent is proportional to the Lyapunov exponent for the blinking vortex flow.

Allshouse and Thiffeault [36] also introduce a topological method for detecting material coherence from a small set of particle trajectories. Coherent regions contain particles that possibly mix with other particles within the region itself but do not mix with particles outside the region. Allshouse and Thiffeault call the set of trajectories arising from particles in a coherent region a coherent bundle. They find coherent bundles by measuring the growth

rate of a topological loop enclosing a pair (or more) of particles (viewed as punctures). Sets of particles enclosed by loops with negligible growth are considered coherent.

2.2 Contributions

As a complement to Thiffeault's work, we analyze the eigenvectors of the Burau representation of a braid of particle trajectories to identify distinct dynamical regimes. We provide a computationally scalable and efficient method that detects coherent sets as levelsets of an eigenvector.

We begin with the lemma that a reducible braid $\alpha \in B_n$ that preserves a family $\mathscr{C}$ of round curves can be written as a product of tubular braids with trivial interior braiding, followed by a product of interior braids with trivial braiding between tubular braids. Each tube is traced out by the path of a simple closed curve in $\mathscr{C}$.

Next, we show that the Burau matrix $B(\alpha)(t)$ of the reducible braid α acts blockwise on a piecewise-constant vector v that respects the structure of α. In particular, we show that if v is constant on the interior of every curve in $\mathscr{C}$, then the image $B(\alpha)(t)v$ is a piecewise-constant vector, constant on the interior of every curve in $\alpha(\mathscr{C}) = \mathscr{C}$. Furthermore, if α is additionally a pure braid, then we can show that $B(\alpha)(t)$ has an eigenvector that is piecewise-constant on its components. Now, every reducible braid is conjugate to a braid that preserves a family of round curves [37–39]. Since $B(\alpha)(t = 1)$ is a permutation matrix, and $B(\alpha)(t)$ is continuous in t, it then follows that for t close to 1, the matrix $B(\alpha)(t)$ has an almost piecewise-constant eigenvector, which is our main result.

We apply our methods to study Aref's blinking vortex flow [28] and a modified Duffing oscillator, which was studied by Allshouse and Thiffeault [36] to test and illustrate the detection of coherent bundles. In the analysis of Aref's blinking vortex flow, our method distinguishes chaotic regions from KAM surfaces, over a range of flow strengths. Meanwhile, in our analysis of the modified Duffing oscillator, in addition to detecting the two dominant regions of mixing detected by Allshouse and Thiffeault, we also detect two additional limit cycles. In these examples, we demonstrate that our methods apply to a broad set of systems, including aperiodic flows and

compressible flows, as long as particle positions do not coincide during the time interval studied. We discuss preprocessing techniques and additional considerations in Section 2.4.2.

Our approach based on topological braids is especially advantageous when data are sparse, since it does not require nearby trajectories or derivatives of the velocity field. However, the braid approach is not without limitations. Accuracy is limited by the length of trajectory histories, and the analysis requires the identification of a projection line upon which distinct trajectories do not coincide.

Throughout this chapter, we assume S is a surface given by the connected sum of $g \geq 0$ tori, with $b \geq 0$ disjoint open disks removed, and $n \geq 0$ punctures.
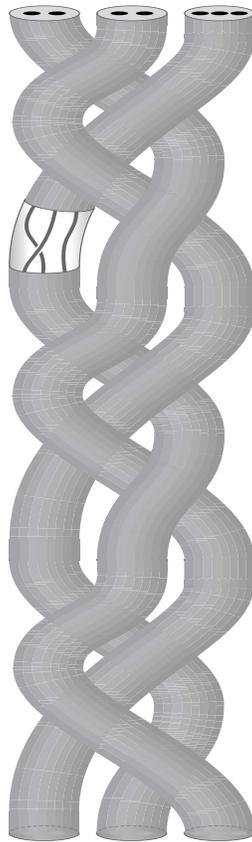
Figure 2.1: Illustration of a tubular braid with non-trivial braiding within a tube. This braid is equivalent to a product of tubular braids with trivial interior braiding, followed by a product of interior braids with trivial braiding between tubular braids.

2.3 Braid groups

There are many equivalent ways of defining the braid group. Each point of view highlights a particular set of characteristics. In this section, we give several definitions of the braid group, so that we can later switch among these viewpoints as convenient.

2.3.1 Braided strands

We begin by describing mathematical braids as a collection of strands from both a geometric and topological point of view, culminating in Artin's classical definition of a braid.

Definition 2.1 (geometric braid [40]). Let $p_1, \cdots, p_n$ be distinguished points in $\mathbb{R}^2$. A *braid* is a collection $\{x_i\}_{i=1}^n$ of n paths $x_i : [0,1] \rightarrow \mathbb{R}^2 \times [0,1]$, $1 \leq i \leq n$, called *strands*, and a permutation $\overline{x}$ of $\{1, \cdots, n\}$ such that each of the following holds:

1. The strands $x_i([0,1])$ are disjoint;

2. $x_i(0) = p_i$;

3. $x_i(1) = p_{\overline{x}(i)}$;

4. $x_i(s) \in \mathbb{R}^2 \times \{s\}$.

A subcollection $\{x_{i_j}\}$ of $\{x_i\}$ is called a *subbraid* of $\{x_i\}$ if $\{x_{i_j}\}$ is also a braid.

A set of n disjoint particle trajectories $x_i : [0,1] \rightarrow X \times [0,1]$, $1 \leq i \leq n$, on a two-dimensional domain $X \subseteq \mathbb{R}^2$ forms a geometric braid on n strands when the final positions $x_i(1)$ are given by a permutation of the of initial positions $x_i(0)$, in which case, each particle trajectory x_i is a strand in space-time $X \times [0,1]$. When the permutation is the identity, we say that the braid is a *pure* braid.

We encode a geometric braid as a *topological braid* by projecting the collection of strands to a fixed plane $\mathbb{R} \times [0,1]$ while retaining information about how strands pass over one another.[†] A *crossing* occurs whenever one strand

---

[†]The choice of projection line is equivalent to a choice of coordinate system. Thus, if the set of particle trajectories forms a geometric braid, then a change in the projection line

passes in front of or behind another. Any geometric braid can be isotoped (i.e., deformed through a continuous family of homeomorphisms) so that at most one crossing occurs at each horizontal level (i.e., at each value of $s \in [0, 1]$). Thus, from each geometric braid, we can specify a topological braid by recording the sequence of crossings.

We enumerate the strands $i = 1, \cdots, n$ according to their ordering on the projection line. As particles move in time, and their strands cross one another, the projection of the strands onto the projection line will change; we update the enumeration accordingly. For $1 \leq i < n$, we let σ_i denote the braid consisting of a single crossing given by passing the ith strand behind the $(i + 1)$st strand (Figure 2.2). Viewed from above, the braid σ_i corresponds to a clockwise half-twist of strands i and $i + 1$. Conversely, we let σ_1^{-1} denote the single crossing given by passing the ith strand in front of the $(i + 1)$st strand.
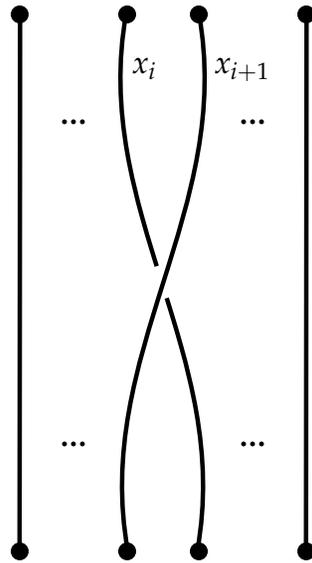
Figure 2.2: Illustration of σ_i, the braid consisting of a single crossing given by passing the ith strand behind x_i the $(i + 1)$st strand x_{i+1}. We adopt the convention of drawing our braids from top to bottom.

A topological braid is specified by a concatenation (product) of σ_i. Following the standard practice in braid literature, we adopt the convention

---

changes the topological braid by conjugation. However, if the set of particle trajectories do not form a geometric braid, then a change of projection line is not guaranteed to merely affect the topological braid by conjugation only [41–43].

of composing our elements from left to right. Furthermore, we adopt the convention of drawing our braids from top to bottom.

Topological braids are classified up to *isotopy*. We say that two braids are *isotopic* if one can be deformed into the other through a continuous family of homeomorphisms.

Definition 2.2 (braid group on n strands). The isotopy classes of braids on n strands form a group B_n, called the *braid group on n strands*.

The collection of braids σ_i, $1 \leq i < n$, generate the braid group B_n. In fact, Artin [44] showed that the braid group B_n admits the following presentation:

$$B_n = \left\langle \sigma_1, \cdots, \sigma_{n-1} \middle| \begin{array}{ll} \sigma_i\sigma_{i+1}\sigma_i = \sigma_{i+1}\sigma_i\sigma_{i+1} & \text{for all } i \\ \sigma_i\sigma_j = \sigma_j\sigma_i & \text{for all } |i-j| > 1 \end{array} \right\rangle.$$

Artin's presentation gives the braid group an algebraic structure. Thus, a product of σ_i will sometimes be referred to as an *algebraic braid*.

For the remainder of this document, when we say braid, we will mean topological braid, unless otherwise specified.

2.3.2 Mapping class groups

Identifying a braid on n strands with (an isotopy class of) a homeomorphism of an n-punctured disk confers several advantages.

First, it is intuitive to visualize the advection of an isotopy class c of a simple closed curve by a homeomorphism of an n-punctured disk. An isotopy class c that encloses two or more punctures may stretch or fold, depending on the dynamics of the homeomorphism. This provides a topological parallel to the advection of material lines in a fluid. Isotopy classes that do not grow (or isotopy classes that grow least [36]) delineate coherent sets.

More importantly, identifying a braid on n strands with a homeomorphism of an n-punctured disk allows us to apply the Nielsen-Thurston classification theorem for surface homeomorphism to obtain a classification of braids. In pursuit of tools for detecting material coherence, we study a particular class of braids given by the Nielsen-Thurston classification system, called *reducible* braids.

We will also note that the homeomorphism viewpoint will also be an important tool in several of our proofs.

Definition 2.3 (mapping class group)**.** The *mapping class group of a surface S*, denoted

$$\mathrm{Mod}(S) = \pi_0\left(\mathrm{Homeo}^+\left(S, \partial S\right)\right),$$

is the group of isotopy classes of elements of $\mathrm{Homeo}^+\left(S, \partial S\right)$, the group of orientation-preserving homeomorphisms on S that restrict to the identity on ∂S, endowed with the compact-open topology.

These groups are discussed in detail by Farb and Margalit [40].

The braid group B_n is isomorphic to the mapping class group of an n-punctured disk D_n:

$$B_n \approx \mathrm{Mod}\left(D_n\right) = \pi_0\left(\mathrm{Homeo}^+\left(D_n, \partial D_n\right)\right). \tag{2.1}$$

Under this isomorphism, each generator σ_i corresponds to the homotopy class of a homeomorphism of D_n that has support a twice-punctured disk and is given by a positive half-twist on this support [40, 45].

Unless otherwise specified, we will use the model of D_n given by the unit disk centered at the origin, with n punctures arranged along the x-axis, so that they partition the interval $[-1, 1]$ into $n + 1$ equal segments.

Nielsen-Thurston classification

The Nielsen-Thurston classification theorem characterizes mapping classes.

Theorem 2.4 (Nielsen-Thurston classification [40, 46])**.** *Let* $g, n \geq 0$. *Each mapping class* $f \in \mathrm{Mod}(S_{g,n})$ *is either periodic, reducible, or pseudo-Anosov. Further, pseudo-Anosov mapping classes are neither periodic nor reducible.*

Our work focuses on reducible mapping classes. Informally, if a mapping class is reducible, we can cut along the simple closed curves of a reduction system and re-apply the theorem to each component of the cut surface. Repeating the process as necessary, we can eventually obtain the result that very mapping class has a representative that decomposes into finite order pieces and pseudo-Anosov pieces. We describe this machinery in detail below.

Definition 2.5 (essential curve). A closed curve is called *essential* if it is not homotopic to a point, a puncture, or a boundary component.

Definition 2.6 (geometric intersection number). The *geometric intersection number* between free homotopy classes *a* and *b* of simple closed curves in a surface *S* is the minimal number of intersection points between a representative curve in the class *a* and a representative curve in the class *b*.

Remark 2.7 ([40]). The geometric intersection number $i(a, b)$ is realized by geodesic representatives of *a* and *b*.

Definition 2.8 (reducible). A mapping class $f \in \text{Mod}(S)$ is *reducible* if there is a non-empty set $\mathscr{C} = \{c_1, \cdots, c_k\}$ of isotopy classes of essential simple closed curves in *S* so that the geometric intersection number $i(c_i, c_j) = 0$ for all *i* and *j* and so that $\{f(c_j)\} = \{c_j\}$. The collection $\mathscr{C}$ is called a *reduction system* for *f*.

We are particularly interested in reducible maps defined on D_n. While the reduction system of a map on D_n may be complicated, every reducible map is conjugate to a braid with a reduction system consisting of a family of geometric ellipses centered on the axis through the punctures [37–39]. We say that a curve is *round* if it is isotopic to a geometric ellipse in D_n. We say that a reduction system $\mathscr{C}$ is *round*[‡] if each curve in $\mathscr{C}$ is round.

In general, a reducible map may admit many reduction systems. A reduction system $\mathscr{C}$ is an *adequate reduction system* for a mapping class $f \in \text{Mod}(S)$ if cutting *S* along $\mathscr{C}$ decomposes *f* so that the restriction of *f* to each component of $S \setminus \mathscr{C}$ is either finite-order or pseudo-Anosov [48].

Birman, Lubotzky, and McCarthy [48] showed that every reducible map *f* has an *essential reduction system*, denoted ERS(*f*), such that

1. every curve[§] $c \in \text{ERS}(f)$ is preserved by some power of *f*;

[‡]Lee and Lee [47] refer to such a reduction system as a *standard reduction system*.

[§]At this point, it should be clear that many characteristics we care about (e.g., braids/mapping classes, reducibility, reduction systems) are given up to isotopy. Thus, for the remainder of our discussion, we will sometimes simply say curve when we mean the isotopy class of a curve.

2. any curve that has non-trivial geometric intersection with some curve $c \in \text{ERS}(f)$ is not preserved by any power of f.

The essential reduction system of a map f is unique (up to isotopy). Furthermore, the essential reduction system of f is a minimal adequate reduction system of f, with respect to inclusion.

The essential reduction system for a mapping class $f \in \text{Mod}(S)$ is sometimes called a *canonical reduction system*, which is defined by Farb and Margalit [40] as the intersection of all maximal (with respect to inclusion) reduction systems for f. The two definitions are equivalent [40, 48].

Remark 2.9 ([39]). With the isomorphism (2.1) in mind, we will sometimes use the term braid to refer to a mapping class.

We say that a braid β is *periodic* if for some integer r, the rth power β^r is isotopic to a full Dehn twist on the boundary of the n-punctured disk D_n.

A braid is *reducible* if its corresponding mapping class is reducible. Similarly, a braid is *pseudo-Anosov* if its corresponding mapping class is pseudo-Anosov.

Remark 2.10. An essential reduction system $\text{ERS}(\beta)$ is non-empty if and only if β is reducible and non-periodic [48].

2.3.3   Automorphisms of a free group

In order to provide an interpretation of the Burau representation as a covering space action in the next section, we use the faithful representation of the braid group B_n as a group of (right) automorphisms of a free group $F_n = \langle x_1, \cdots, x_n \rangle$ on n generators [44, 49].

The representation $\xi : B_n \to \text{Aut}(F_n)$ is induced by mapping a generator σ_i of the braid group B_n to the automorphism $(\sigma_i)\xi$ of a free group F_n, given by

$$x_i \to x_i x_{i+1} x_i^{-1},$$
$$x_{i+1} \to x_i,$$
$$x_j \to x_j \quad \text{if } j \neq i, i+1.$$

Remark 2.11. The automorphism $(\sigma_i^{-1})\xi$ is given by

$$x_i \rightarrow x_{i+1},$$
$$x_{i+1} \rightarrow x_{i+1}^{-1} x_i x_{i+1},$$
$$x_j \rightarrow x_j \quad \text{if } j \neq i, i+1.$$

2.4 Application to the analysis of flows

In this section, we detect dynamically distinct regions as levelsets of an eigenvector of the (unreduced) Burau matrix corresponding to the motion of the particles. To do so, we introduce the Burau representation, which is an example of a Magnus representation [44]. We interpret the Burau representation as a covering space action. Following this, we show that when a reducible pure braid α has a round reduction system $\mathscr{C}$, the Burau matrix $B(\alpha)(t)$ has an eigenvector that is piecewise constant on the interior of the reduction curves $c_j \in \mathscr{C}$. We conclude that for any reducible braid β, the Burau matrix $B(\beta)(t \approx 1)$ has an eigenvector that is approximately piecewise-constant on components of its Nielsen-Thurston decomposition.

2.4.1 (Unreduced) Burau representation

Matrix representation

The Burau representation is a homomorphism

$$\rho_n : B_n \rightarrow \mathrm{GL}_n(\Lambda),$$

with $\Lambda := \mathbb{Z}[t^{\pm}]$. The matrix representation allow us to examine the dynamics of a motion of particles through a study of eigenvectors. There are two forms of the Burau representation, the reduced and unreduced. Our analysis focuses on the unreduced Burau representation, which is given by mapping a generator σ_i of B_n to the block matrix

$$B(\sigma_i)(t) := I_{i-1} \oplus \begin{pmatrix} 1-t & t \\ 1 & 0 \end{pmatrix} \oplus I_{n-i-1}. \tag{2.2}$$

We write $B(\alpha)(t)$ for the Burau matrix for the braid α with parameter t.

We note that

$$[B(\sigma_i)(t)]^{-1} = I_{i-1} \oplus \begin{pmatrix} 0 & 1 \\ t^{-1} & 1-t^{-1} \end{pmatrix} \oplus I_{n-i-1}.$$

It can be verified that

$$[B\left(\sigma_i\right)(t)]^{-1} = B\left(\sigma_i^{-1}\right)(t).$$

The constant vector $(1, \cdots, 1)$ generates an invariant subspace of the matrices $B\left(\sigma_i\right)(t)$. The *reduced Burau representation* is obtained by taking the quotient.

Unless otherwise specified, in this document, when we say Burau representation, we will mean the unreduced Burau representation.

The Burau representation is known to be not faithful for $n \geq 5$ [50, 51]. The case $n = 4$ remains open. Church and Farb [52] that the kernel of ρ_n is in fact quite large for $n \geq 6$. In our experience, the unfaithfulness of the Burau representation has not been an impediment to extracting coherent sets from physical systems.

Remark 2.12. For every constant vector v, we have that $B(\sigma)(t)v = v$ for every braid σ and all t.

Proof.

It is sufficient to prove this result for the generators $B(\sigma_i)(t)$.

When $v_1 = v_2$, we have that

$$\begin{pmatrix} 1-t & t \\ 1 & 0 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} (1-t)v_1 + tv_2 \\ v_1 \end{pmatrix} = \begin{pmatrix} v_1 \\ v_1 \end{pmatrix}.$$

Thus,

$$B\left(\sigma_i\right)(t) := I_{i-1} \oplus \begin{pmatrix} 1-t & t \\ 1 & 0 \end{pmatrix} \oplus I_{n-i-1}$$

fixes constant vectors.   □

Covering space action

The Burau representation has a topological interpretation as a covering space action [53, 54].

For $1 \leq j \leq n$, let $p_j = \frac{j}{n+1}$, and consider an n-punctured disk

$$D_n = \left\{\left|z - \frac{1}{2}\right| \leq 1\right\} \setminus \{p_1, \cdots, p_n\}.$$

Fix a basepoint $p_0 \in \partial D_n$, and around each puncture p_j, take a small clockwise loop $\mathbf{x}_j$, such that $\mathbf{x}_j(0) = \mathbf{x}_j(1) = p_0$. Then $\pi_1(D_n, p_0)$ is freely generated by the set of homotopy classes $[\mathbf{x}_j]$.

Let $\tau : \pi_1(D_n, p_0) \to \mathbb{Z}$ be the epimorphism generated by $\tau([\mathbf{x}_j]) = 1$ for $1 \le j \le n$. The kernel of τ consists of all words in $\{\mathbf{x}_j\}_{j=1}^n$ whose exponent sum is zero. This is a normal subgroup. So we can define a covering space $p : \widetilde{D}_n \to D_n$ corresponding to the kernel of τ.

Let F be the fiber above the basepoint p_0, and consider the relative homology group $H_1(\widetilde{D}_n, F)$. By construction, the deck group of $\widetilde{D}_n$ is isomorphic to $\mathbb{Z}$. Call its generator t. Then $H_1(\widetilde{D}_n, F)$ is free and n-dimensional as a module over $\Lambda = \mathbb{Z}[t^{\pm}]$. The action of B_n on $H_1(\widetilde{D}_n, F)$ gives the (unreduced) Burau representation.

Example 2.13. We consider the case $n = 3$.

Let D_3 denote a thrice-punctured disk with punctures p_1, p_2, p_3. Fix a basepoint $p_0 \in \partial D_3$. Let $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$ be small clockwise loops in (D_3, p_0), enclosing punctures p_1, p_2, p_3, respectively.

As a Λ-module, the homology group $H_1(\widetilde{D}_3, F)$ has three generators, $\tilde{\mathbf{x}}_1$, $\tilde{\mathbf{x}}_2$, $\tilde{\mathbf{x}}_3$, given by the lifts of $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$, each $\tilde{\mathbf{x}}_i$ beginning at some fixed basepoint $\tilde{p}_0 \in F$.

The generator σ_1 induces a map on the fundamental group $\pi_1(D_3, p_0)$ sending $\mathbf{x}_1$ to $\mathbf{x}_1 \mathbf{x}_2 \mathbf{x}_1^{-1}$. Thus, σ_1 induces a map $(\sigma_1)_*$ on homology that sends $\tilde{\mathbf{x}}_1$ to the lift of $\mathbf{x}_1 \mathbf{x}_2 \mathbf{x}_1^{-1}$, which is

$$\tilde{\mathbf{x}}_1 + t\tilde{\mathbf{x}}_2 - t\tilde{\mathbf{x}}_1 = (1-t)\tilde{\mathbf{x}}_1 + t\tilde{\mathbf{x}}_2.$$

Similarly,

$$(\sigma_1)_*(\tilde{\mathbf{x}}_2) = \tilde{\mathbf{x}}_1,$$
$$(\sigma_1)_*(\tilde{\mathbf{x}}_3) = \tilde{\mathbf{x}}_3.$$

Let B_1 be the matrix whose (i, j)th entry is the coefficient of $\tilde{\mathbf{x}}_j$ in the image $(\sigma_1)_*(\tilde{\mathbf{x}}_i)$. Then

$$B_1 = \begin{pmatrix} 1-t & t & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = B(\sigma_1)(t).$$

Performing the same analysis for the generator σ_2, we have

$$(\sigma_2)_*(\tilde{\mathbf{x}}_1) = \tilde{\mathbf{x}}_1,$$
$$(\sigma_2)_*(\tilde{\mathbf{x}}_2) = (1 - t)\tilde{\mathbf{x}}_2 + \tilde{\mathbf{x}}_3,$$
$$(\sigma_2)_*(\tilde{\mathbf{x}}_3) = \tilde{\mathbf{x}}_2.$$

As before, we let B_2 be the matrix whose (i, j)th entry is the coefficient of $\tilde{\mathbf{x}}_j$ in the image of $(\sigma_2)_*(\tilde{\mathbf{x}}_i)$. Then

$$B_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1-t & t \\ 0 & 1 & 0 \end{pmatrix} = B(\sigma_2)(t).$$

Relationship to algebraic intersection numbers

The matrix entries of the Burau representation can also be interpreted as algebraic intersection numbers [50, 55].

Definition 2.14 (algebraic intersection number [40]). Let a and b be a pair of transverse, oriented, simple closed curves in a surface S. The *algebraic intersection number* $\hat{i}(a, b)$ is defined as the sum of the indices of the intersection points of a and b, where an intersection point is of index $+1$ when the orientation of the intersection agrees with the intersection of S and -1 otherwise.

Remark 2.15 ([40]). The algebraic intersection number $\hat{i}(a, b)$ depends only on the homology classes of a and b.

Let μ_j, $1 \le j \le n$, be the arcs shown in Figure 2.3.

We define a map $\int \omega_j : H_1(\widetilde{D}_n, F) \to \Lambda$ on a homology class $a \in H_1(\widetilde{D}_n, F)$ by setting

$$\int_a \omega_j = \sum_{k \in \mathbb{Z}} t^k \, \hat{i}(a, t^k \mu_j),$$

where $\hat{i}(a, t^k \mu_j)$ is the algebraic intersection number of the arcs a and $t^k \mu_j$ in the covering space $\widetilde{D}_n$.

Then the Burau matrix $B(\beta)$ of a braid $\beta \in B_n$ has entries given by

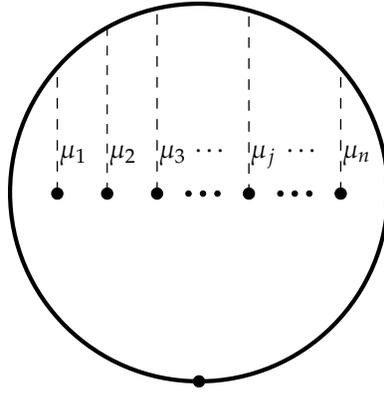$$B(\beta)(t)_{i,j} = \int_{\beta(\mathbf{x}_i)} \omega_j.$$

Figure 2.3: Illustration of μ_j in $\widetilde{D}_n$.

Example 2.16. In Figures 2.4 and 2.5, we consider the image of the clockwise loops $\mathbf{x}_i$ and $\mathbf{x}_{i+1}$, respectively, under the braid σ_i. We consider the lifts of $\sigma_i(\mathbf{x}_i)$ and $\sigma_i(\mathbf{x}_{i+1})$ to the cyclic covering space $\widetilde{D}_n$. The oriented intersections of the lifts with the arcs $\mu_i, t^1\mu_i, t^2\mu_i$ and $\mu_{i+1}, t^1\mu_{i+1}, t^2\mu_{i+1}$ (illustrated by the solid lines in each copy of D_n in $\widetilde{D}_n$) give the matrix entries for the Burau matrix $B(\sigma_i)(t)$ in Equation (2.2).

In particular, the lift of $\sigma_i(\mathbf{x}_i)$ begins at $\widetilde{p_0}$, intersects μ_i once, then intersects $t\mu_{i+1}$ once, ascends to the next copy of D_n before descending again, upon which it intersects $t\mu_i$ in the opposite orientation, and ends at $t\widetilde{p_0}$. (See Figure 2.4 for an illustration.) This corresponds to the ith row of the Burau matrix $B(\sigma_i)(t)$, which has non-trivial entries $1 - t$ and t in columns i and $i + 1$, respectively.

The lift of $\sigma_i(\mathbf{x}_{i+1})$ begins at $\widetilde{p_0}$, intersects μ_i once, and ends at $t\widetilde{p_0}$. (See Figure 2.5 for an illustration.) This corresponds to the $(i + 1)$st row of the Burau matrix $B(\sigma_i)(t)$, which has a single non-trivial entry consisting of 1 in column i.

Example 2.17 (tubular braid). We use the notation of Band and Boyland [56] and refer to the braid that moves the group of η_1 consecutive strands starting at strand i behind the group of η_2 consecutive strands starting at strand $i + \eta_1$ as

$$\sigma_{i,\eta_1,\eta_2} = \left(\sigma_{i+\eta_1-1} \cdots \sigma_{i+\eta_1+\eta_2-2}\right)\left(\sigma_{i+\eta_1-2} \cdots \sigma_{i+\eta_1+\eta_2-3}\right) \cdots \left(\sigma_i \cdots \sigma_{i+\eta_2-1}\right).$$

In particular, $\sigma_{i,1,1} = \sigma_i$ for all $1 \le i < n$.
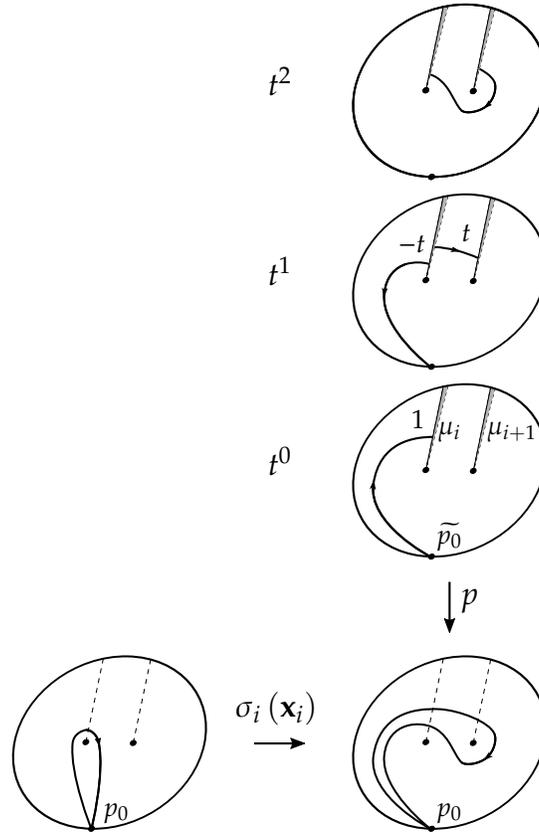
Figure 2.4: Illustration of $\sigma_i(\mathbf{x}_i)$ and its lift to the covering space $p : \widetilde{D}_n \to D_n$.

By considering lifts of the images of the loops $\mathbf{x}_j$ under σ_{i,η_1,η_2} to the cyclic cover $\widetilde{D}_n$ (see Figure 2.7), one can prove that the Burau matrix $B(\sigma_{i,\eta_1,\eta_2})$ of the braid σ_{i,η_1,η_2} is a block matrix of the form:

$$
I_{i-1} \oplus \left(
\begin{array}{cccc|ccc}
1-t & t-t^2 & \cdots & t^{\eta_2-1}-t^{\eta_2} & t^{\eta_2} & \cdots & 0 \\
\vdots & \vdots & & \vdots & & \ddots & \\
1-t & t-t^2 & \cdots & t^{\eta_2-1}-t^{\eta_2} & 0_{\eta_1} & \cdots & t^{\eta_2} \\
\hline
& I_{\eta_2} & & & & 0 &
\end{array}
\right) \oplus I_{n-i-\eta_1-\eta_2+1}.
$$

Given a block vector $v = (v_1, \cdots, v_n)$, with

$$
v_j = \begin{cases} b & \text{for } i \leq j < i+\eta_1, \\ c & \text{for } i+\eta_1 \leq j < i+\eta_1+\eta_2, \end{cases}
$$
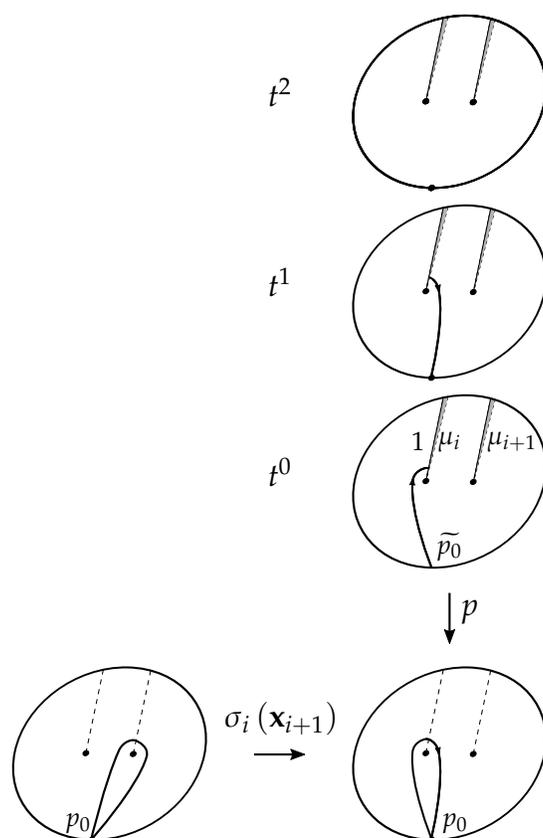
Figure 2.5: Illustration of $\sigma_i(\mathbf{x}_{i+1})$ and its lift to the covering space $p : \widetilde{D}_n \to D_n$.

we have that

$$
B\left(\sigma_{i,\eta_1,\eta_2}\right)(t)
\begin{pmatrix}
v_1 \\
\vdots \\
v_{i-1} \\
b \\
\vdots \\
b \\
c \\
\vdots \\
c \\
v_{i+\eta_1+\eta_2} \\
\vdots \\
v_n
\end{pmatrix}
=
\begin{pmatrix}
v_1 \\
\vdots \\
v_{i-1} \\
b\left(1-t^{\eta_2}\right)+ct^{\eta_2} \\
\vdots \\
b\left(1-t^{\eta_2}\right)+ct^{\eta_2} \\
b \\
\vdots \\
b \\
v_{i+\eta_1+\eta_2} \\
\vdots \\
v_n
\end{pmatrix}.
$$

Figure 2.6: Illustration of a braid corresponding to σ_{i,η_1,η_2}

So when t is an η_2th root of unity, the Burau matrix $B(\sigma)(t)$ is a permutation of the blocks indexed by $i \leq \mu < i + \eta_1$ and $i + \eta_1 \leq \nu < i + \eta_1 + \eta_2$ and the identity elsewhere. For all other t, the Burau matrix $B(\sigma)(t)$ maps block vectors v that are piecewise-constant on the blocks indexed by $i \leq \mu < i + \eta_1$ and $i + \eta_1 \leq \nu < i + \eta_1 + \eta_2$ to block vectors that are piecewise-constant on the blocks indexed by indexed by $i \leq \nu < i + \eta_2$ and $i + \eta_1 \leq \mu < i + \eta_1 + \eta_2$.

Performing a similar analysis, one can show that the Burau matrix $B(\sigma_{i,\eta_1,\eta_2}^{-1})(t)$ of the inverse $\sigma_{i,\eta_1,\eta_2}^{-1}$ is a block matrix of the form

$$
I_{i-1} \oplus \left(\begin{array}{ccc|cccc}
\multicolumn{3}{c|}{0_{\eta_2}} & \multicolumn{4}{c}{I_{\eta_1}} \\
t^{-\eta_1} & \cdots & 0 & 1 - t^{-1} & t^{-1} - t^{-2} & \cdots & t^{-\eta_1-1} - t^{\eta_1} \\
& \ddots & & \vdots & \vdots & & \vdots \\
0 & \cdots & t^{-\eta_1} & 1 - t^{-1} & t^{-1} - t^{-2} & \cdots & t^{-\eta_1-1} - t^{\eta_1}
\end{array} \right) \oplus I_{n-1-\eta_1-\eta_2+1},
$$

Figure 2.7: **Covering space action of** σ_{i,η_1,η_2}. If $i \leq j < i + \eta_1$ (top two rows), the lift of $\sigma_{i,\eta_1,\eta_2}(\mathbf{x}_j)$ to $\widetilde{D}_n$ begins at $\widetilde{p}_0$ and goes up, intersecting the arcs $\mu_i, t\mu_{i+1}, \cdots, t^{\eta_2-1}\mu_{i+\eta_2}$; it then intersects $t^{\eta_2}\mu_{\eta_2+j-i+1}$ before going back down, upon which it intersects the arcs $t^{\eta_2}\mu_{i+\eta_2}, t^{\eta_2-1}\mu_{i+\eta_2-1}, \cdots, t\mu_i$ in the opposite direction, and ends at $t\widetilde{p}_0$. If $i + \eta_1 \leq j < i + \eta_1 + \eta_2$ (bottom two rows), the lift of $\sigma_{i,\eta_1,\eta_2}(\mathbf{x}_j)$ begins at $\widetilde{p}_0$, intersects $\mu_{j-\eta_1}$, and ends on $t\widetilde{p}_0$.

and

$$B\left(\sigma_{i,\eta_1,\eta_2}^{-1}\right)(t) \begin{pmatrix} v_1 \\ \vdots \\ v_{i-1} \\ b \\ \vdots \\ b \\ c \\ \vdots \\ c \\ v_{i+\eta_1+\eta_2} \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} v_1 \\ \vdots \\ v_{i-1} \\ c \\ \vdots \\ c \\ bt^{-\eta_1} + c\left(1 - t^{-\eta_1}\right) \\ \vdots \\ bt^{-\eta_1} + c\left(1 - t^{-\eta_1}\right) \\ v_{i+\eta_1+\eta_2} \\ \vdots \\ v_n \end{pmatrix}.$$

Example 2.17 is an example of a *tubular braid* [57].

Decomposition of reducible braids

Next, we show that a reducible braid with a round reduction system can be decomposed as a product of braiding between tubular braids $\sigma_{i(\ell),\eta_1(\ell),\eta_2(\ell)}^{\varepsilon(\ell)}$ with trivial braiding within tubes, where $i(\ell), \eta_1(\ell), \eta_2(\ell)$ are positive integers and $\varepsilon(\ell) = \pm 1$, followed by a product of braiding within tubular braids with trivial braiding between tubes.

Note that given any reduction system $\mathscr{C}$ of a reducible braid $\alpha \in B_n$, we can always choose a non-empty, non-nested subset $\mathscr{C}_{\text{ext}}$ consisting only of the outermost (i.e., exterior) curves of $\mathscr{C}$.

Lemma 2.18. *Let $\alpha \in B_n$ be a reducible braid that preserves a non-nested round reduction system $\mathscr{C}$. For each puncture p_r of D_n not enclosed by any curve of $\mathscr{C}$, extend $\mathscr{C}$ by adding a round curve c_r enclosing the single puncture p_r. Let k be the number of curves in (the newly extended) $\mathscr{C}$. Then there exists a finite sequence of tuples $(i(\ell), \eta_1(\ell), \eta_2(\ell), \varepsilon(\ell))$ of integers and a finite sequence of braids α_j, $1 \le j \le k$, each α_j supported on the punctured disk with boundary given by the curve $c_j \in \mathscr{C}$, such that*

$$\alpha = \prod_\ell \sigma_{i(\ell),\eta_1(\ell),\eta_2(\ell)}^{\varepsilon(\ell)} \cdot \prod_{j=1}^k \alpha_j. \tag{2.3}$$

Proof.

Let E_j denote the punctured disk enclosed by the circle $c_j \in \mathscr{C}$, enumerating the disks E_j in order along the axis through the punctures. Let

$$\widehat{D}_k = D_n \setminus \bigcup_{j=1}^{k} E_j.$$

Then α induces an automorphism $\widehat{\alpha}$ on $\widehat{D}_k$.

If we collapse each hole of $\widehat{D}_k$ to a puncture, we can regard $\widehat{D}_k$ as a k-punctured disk. So the braid $\widehat{\alpha}$ can be given by a sequence of Artin generators for the braid group B_k:

$$\widehat{\alpha} = \prod \widehat{\sigma}_{a_\ell}^{\varepsilon_\ell},$$

where $1 \leq a_\ell < k$ and $\varepsilon_\ell = \pm 1$ for all ℓ. Viewed from above, each generator $\widehat{\sigma}_{a_\ell}$ corresponds to a clockwise half-twist interchanging the holes obtained by removing E_{a_ℓ} and $E_{a_\ell+1}$. Thus, each $\widehat{\sigma}_{a_\ell}^{\varepsilon_\ell}$ specifies a tuple $(i(\ell), \eta_1(\ell), \eta_2(\ell), \varepsilon(\ell))$, given by the minimum index of the punctures in E_{a_ℓ}, the number of punctures in E_{a_ℓ}, the number of punctures in $E_{a_\ell+1}$, and the direction of the half-twist, respectively.

Letting $\alpha_j = \alpha \mid_{E_j}$ denote the restriction of α to E_j, this completes the decomposition.

$\square$

Definition 2.19 (piecewise-constant vector). Let $\mathscr{C}$ be a non-empty finite collection of pairwise-disjoint non-trivial simple closed curves such that each curve $c_j \in \mathscr{C}$ encloses at least one puncture of an n-punctured disk D_n. Cutting the n-punctured disk D_n along the curves $c_j \in \mathscr{C}$, we obtain a collection $\mathscr{E}$ of path-components. We say that a vector $v = (v_1, \cdots, v_n)$ is *piecewise-constant on components of* $(D_n, \mathscr{C})$ if $v_\ell = v_{\ell'}$ whenever the corresponding punctures p_ℓ and $p_{\ell'}$ belong to the same path-component in $\mathscr{E}$.

Let E_j denote the punctured disk enclosed by the curve $c_j \in \mathscr{C}$. We say that the vector v is *constant on* E_j if $v_\ell = v_{\ell'}$ for all punctures $p_\ell, p_{\ell'}$ in E_j.

Example 2.20. Let $\mathscr{C} = \{c_1, c_2, c_3, c_4\}$ be a finite collection of pairwise-disjoint simple closed curves such that:

- the curve c_1 encloses punctures $p_1, \cdots, p_{i-1}$,

- the curve c_2 encloses punctures $p_i, \cdots, p_{i+\eta_1-1}$,

- the curve c_3 encloses punctures $p_{i+\eta_1}, \cdots, p_{i+\eta_1+\eta_2-1}$, and

- the curve c_4 encloses punctures $p_{i+\eta_1+\eta_2}, \cdots, p_n$.

If v be a vector that is piecewise-constant on components of $(D_n, \mathscr{C})$, then by Example 2.17, the image $B(\sigma_{i,\eta_1,\eta_2})(t)v$ is also piecewise-constant on components of $(D_n, \sigma_{i,\eta_1,\eta_2}(\mathscr{C}))$ for all t.

Lemma 2.21. *Consider the braid*

$$\alpha = \prod_{\ell=1}^{m} \sigma_{i(\ell),\eta_1(\ell),\eta_2(\ell)}^{\varepsilon(\ell)} \cdot \prod_{j=1}^{k} \alpha_j,$$

from Equation (2.3), where $(i(\ell), \eta_1(\ell), \eta_2(\ell), \varepsilon(\ell))$ and α_j are given in the proof of Lemma 2.18. Let $\mathscr{C}$ be the non-nested family of round simple closed curves given in the hypothesis of Lemma 2.18. If v be a block vector in $\mathbb{C}^n$ such that v is piecewise-constant on components of $(D_n, \mathscr{C})$, then the image $B(\alpha)(t)v$ is piecewise-constant on components of $(D_n, \alpha(\mathscr{C}))$.

Proof.

As a corollary of Remark 2.12, for each $1 \le j \le k$, we have that $B(\alpha_j)(t)v = v$ for all vectors v that are constant on E_j. So

$$\left(\prod_{j=1}^{k} B(\alpha_j)(t)\right)v = v.$$

Thus, without loss of generality, we need only consider braids of the form $\alpha = \prod_{\ell=1}^{m} \sigma_{i(\ell),\eta_1(\ell),\eta_2(\ell)}$.

We prove the proposition by induction on $m \ge 1$.[¶]

[¶]We note that an intermediate braid $\prod_{\ell=1}^{m'} \sigma_{i(\ell),\eta_1(\ell),\eta_2(\ell)}$, where $1 \le m' < m$, is not necessarily reducible. However, from the proof of Lemma 2.18, we can guarantee that every intermediate braid is a tubular braid, where each tube is delineated by the trajectory of some curve in $\mathscr{C}$. Thus, for the purposes of induction, we do not think of $\mathscr{C}$ as a reduction system but merely a non-nested family of round simple closed curves that decompose α into its tubular structure.

The base case is given by Example 2.17.

For the inductive case, we consider $\alpha = \prod_{\ell=1}^{m} \sigma_{i(\ell),\eta_1(\ell),\eta_2(\ell)}$ and the corresponding Burau matrices

$$B(\alpha)(t) = \prod_{\ell=1}^{m} B\left(\sigma_{i(\ell),\eta_1(\ell),\eta_2(\ell)}\right)(t).$$

Let $\alpha' = \prod_{\ell=1}^{m-1} \sigma_{i(\ell),\eta_1(\ell),\eta_2(\ell)}$. By the inductive hypothesis,

$$w' = B(\sigma')(t)v$$

is piecewise-constant on components of $(D_n, \alpha'(\mathscr{C}))$.

Furthermore, by construction (Lemma 2.18), the image w' is piecewise-constant on the component containing $i(m)$ and on the component containing $i(m) + \eta_1(m)$. So applying the base case (Example 2.17), we have that

$$B(\sigma_{i(m),\eta_1(m),\eta_2(m)})(t)w'$$

is piecewise-constant on components of $(D_n, \alpha(\mathscr{C}))$.

This proves the inductive hypothesis and thus concludes our proof.

$\square$

Piecewise-constant eigenvectors

Combining Lemmata 2.18 and 2.21, we have the following corollary:

Corollary 2.22. *Let α be a reducible braid with a round reduction system $\mathscr{C}$. Let V be a block vector in $\mathbb{C}^n$ that is piecewise-constant on components of $(D_n, \mathscr{C})$. Then the image $B(\alpha)(t)V$ is a block vector, piecewise-constant on components of $(D_n, \mathscr{C}) = (D_n, \alpha(\mathscr{C}))$.*

Suppose the block vector V has value V_j on E_j. Then we can write

$$B(\alpha)(t) \begin{pmatrix} V_1 \\ \vdots \\ V_k \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{k} V_1 p_{1,j}(t) \\ \vdots \\ \sum_{j=1}^{k} V_k p_{k,j}(t) \end{pmatrix},$$

where each $p_{i,j}(t)$ is a polynomial in t.

For each fixed t, the matrix

$$
P = \begin{pmatrix} p_{1,1}(t) & \cdots & p_{1,k}(t) \\ \vdots & \ddots & \vdots \\ p_{k,1}(t) & \cdots & p_{k,k}(t) \end{pmatrix}
$$

has an eigenpair (λ, v), with $\lambda \in \mathbb{C}$ and $v \in \mathbb{C}^k$. Letting the jth block of V be given by the jth entry of v, we have that (λ, V) is an eigenpair of $B(\alpha)(t)$.

Now, when α is additionally a pure braid, the Burau matrix $B(\alpha)(1)$ is the identity. Since the Burau representation is continuous in t at $t = 1$, then for every $\varepsilon > 0$, the Gershgorin circle theorem guarantees that there exists $\delta > 0$ such that for all t satisfying $|t - 1| < \delta$, we have that $|\lambda - 1| < \varepsilon$. (See Lemma 2.26 for guidance on the choice of t.)

Thus,

Proposition 2.23. *Let α be a reducible pure braid with round reduction system $\mathscr{C}$. Then there exists $\delta > 0$ such that for all t satisfying $|t - 1| < \delta$, we have that the Burau matrix $B(\alpha)(t)$ has an eigenvector that is piecewise constant on $(D_n, \mathscr{C})$.*

Using that every reducible braid is conjugate to a braid with a round reduction system, we now show that every reducible pure braid has an eigenvector that is almost piecewise-constant on its components.

Proposition 2.24. *Let $\beta \in B_n$ be a reducible pure braid with a reduction system $\mathscr{C}_\beta$. For each $\varepsilon > 0$, there exists $t \in \mathbb{C}$, such that the Burau matrix $B(\beta)(t)$ has an eigenvector $v_\varepsilon(t) = v(t) + e(t)$, where $v(t)$ is a vector that is piecewise-constant on $(D_n, \mathscr{C}_\beta)$, and $e(t)$ is a matrix whose entries are bounded by ε, with $\left|e_{i,j}(t)\right| < \varepsilon$ for all $1 \leq i, j \leq n$.*

Proof.

There exists a braid α conjugate to β in B_n, with $\beta = \gamma^{-1}\alpha\gamma$ for some $\gamma \in B_n$, such that α has a round reduction system $\mathscr{C} = \gamma(\mathscr{C}_\beta)$, whose curves are given by $\gamma(c_j)$, for some curve $c_j \in \mathscr{C}_\beta$, where we consider γ as an automorphism of the punctured disk as necessary [57].

Let their corresponding Burau matrices be denoted

$$A(t) = B(\alpha)(t),$$
$$B(t) = B(\beta)(t),$$
$$C(t) = B(\gamma)(t).$$

Then $B(t) = C(t)A(t)C(t)^{-1}$. (Note that we compose braids from left to right but compose matrices from right to left.)

Since $C(t)$ is continuous in t at $t = 1$, there exists $\delta_C > 0$ such that for all t satisfying $|t - 1| < \delta_C$, we have that

$$\left\| C(t) - C(1) \right\| < \varepsilon. \tag{2.4}$$

(See Lemma 2.26.)

By Proposition 2.23, since α is a reducible pure braid with a round reduction system $\mathscr{C}$, there exists $\delta_A > 0$ such that for all t satisfying $|t - 1| < \delta_A$, the Burau matrix $A(t)$ has an eigenvector $v(t)$ that is piecewise-constant on $(D_n, \mathscr{C})$.

So for all t within $\delta = \min\{\delta_A, \delta_C\}$ of one, we have that Equation (2.4) holds and that $A(t)$ has an eigenvector $v(t)$ that is piecewise-constant on $(D_n, \mathscr{C})$.

Without loss of generality, assume $\left\| v(t) \right\| = 1$. Note that $C(t)v(t)$ is an eigenvector of $B(t)$,

$$B(t)C(t)v(t) = C(t)A(t)C(t)^{-1}C(t)v(t) = C(t)A(t)v(t) = C(t)v(t),$$

but not necessarily piecewise-constant on $(D_n, \gamma^{-1}(\mathscr{C}))$, where $\gamma^{-1}(\mathscr{C})$ is the reduction system $\mathscr{C}_\beta$ for β. On the other hand, since $C(1)$ is a permutation matrix, the vector $C(1)v(t)$ is piecewise-constant on $(D_n, \gamma^{-1}(\mathscr{C}))$ but not necessarily an eigenvector of $B(t)$.

Since

$$\left\| C(t)v(t) - C(1)v(t) \right\| \le \left\| C(t) - C(1) \right\| \left\| v(t) \right\| = \left\| C(t) - C(1) \right\| < \varepsilon,$$

the vector $C(t)v(t)$ satisfies the proposition.

$$\square$$

Remark 2.25. Since the Burau representation is not faithful, we cannot guarantee that the eigenvector $v(t)$ is non-constant.

Continuity of the Burau representation

Since the Burau representation is continuous in t at $t = 1$, then for every braid γ and every $\varepsilon > 0$, there exists t such that $\left\| B(\gamma)(t) - B(\gamma)(1) \right\| < \varepsilon$. Informally, since the Burau matrix $B(\gamma)(1)$ is a permutation matrix, then for $t \approx 1$, the Burau matrix $B(\gamma)(t)$ is approximately a permutation matrix. In the following lemma, we quantify this approximation.

Lemma 2.26. *Let $\gamma \in B_n$, with $\gamma = \prod_{j=1}^{L} \sigma_{a_j}$. Then for all $0 < |t| < 1$, we have*

$$\left\| B(\gamma)(t) - B(\gamma)(1) \right\| \le L\,(n+1)^{\frac{L-1}{2}}\,\sqrt{2}\,|1 - t|,$$

where $\|\cdot\|$ is the Frobenius matrix norm,

$$\|B\| = \left(\sum_{j=1}^{n} \sum_{i=1}^{n} |b_{i,j}|^2 \right)^{1/2}.$$

Proof.

For each generator $\sigma_i \in B_n$ and $t \in \mathbb{C}$ such that $|t| \le 1$, we have

$$\left\| B(\sigma_i)(t) \right\|^2 = n - 1 + |1 - t|^2 + |t|^2 \le n + 1,$$
$$\left\| B(\sigma_i)(1) \right\|^2 = n,$$

and

$$\left\| B(\sigma_i)(t) - B(\sigma_i)(1) \right\|^2 = 2\,|1 - t|^2.$$

So for a braid $\gamma = \prod_{j=1}^{L} \sigma_{a_j}$, we have that

$$\left\| \prod_{j=1}^{L} B(\sigma_{a_j})(t) - \prod_{j=1}^{L} B(\sigma_{a_j})(1) \right\|,$$

$$= \left\| \sum_{j=1}^{L} \left(\prod_{i=1}^{j-1} B(\sigma_{a_j})(t) \right) \cdot \left(B(\sigma_{a_j})(t) - B(\sigma_{a_j})(1) \right) \cdot \left(\prod_{i=j+1}^{L} B(\sigma_{a_i})(1) \right) \right\|,$$

$$\le \sum_{j=1}^{L} \left\| \left(\prod_{i=1}^{j-1} B(\sigma_{a_j})(t) \right) \cdot \left(B(\sigma_{a_j})(t) - B(\sigma_{a_j})(1) \right) \cdot \left(\prod_{i=j+1}^{L} B(\sigma_{a_i})(1) \right) \right\|,$$

$$\le \sum_{j=1}^{L} \left(\prod_{i=1}^{j-1} \left\| B(\sigma_{a_j})(t) \right\| \right) \cdot \left\| B(\sigma_{a_j})(t) - B(\sigma_{a_j})(1) \right\| \cdot \left(\prod_{i=j+1}^{L} \left\| B(\sigma_{a_i})(1) \right\| \right),$$

$$\le L\,(n+1)^{\frac{L-1}{2}}\,\sqrt{2}\,|1 - t|. \qquad \square$$

Corollary 2.27. *Let $\varepsilon > 0$. For $t \in \mathbb{C}$, with $0 < |t| < 1$, such that*

$$|1 - t| < \frac{\sqrt{2}\varepsilon}{2L}(n+1)^{-\frac{L-1}{2}}.$$

Then $\left\| B(\gamma)(t) - B(\gamma)(1) \right\| < \varepsilon$.

2.4.2 Numerical implementation

In the following section, we compute the Burau matrix for two different dynamical systems and visualize corresponding eigenvectors whose level-sets correspond to components of the Nielsen-Thurston decomposition.

Given trajectories $x_j : [0,1] \to \mathbb{R}^2 \times [0,1]$, $1 \leq j \leq n$, we use braidlab [43] to compute the algebraic braid β corresponding to the motion of the particles, with respect to some fixed projection line. Prior to computing the algebraic braid, some preprocessing and extra consideration are necessary.

(i) **Resolving coincident projections.** We note that when our initial positions are given by grid points of a regular grid, it can be helpful to perturb the initial positions and/or choose a projection line that is not parallel to the grid lines (i.e., not the x- or y-axis). This will help resolve some coincident trajectories in the projection, allowing the algebraic braid (i.e., sequence of σ_i) to be computed.

(ii) **Enforcing closure.** In general, sets of trajectories do not form true geometric braids, since the set of particle positions at initial time and the set of particle positions at final time are not necessarily the same. In such a setting, changing the projection line (to form an algebraic braid) will not necessarily result in a conjugate braid [41, 42]. To rectify this, in the examples that follow, for each trajectory $\left(x_j(t_0), \cdots x_j(t_N)\right)$, we append the initial position $x_j(t_0)$ to form the closed trajectory

$$\left(x_j(t_0), \cdots x_j(t_N), x_j(t_0)\right).$$

We note that there are many ways one may choose to enforce closure. In braidlab, Thiffeault provides a closure method that simply draws line segments from the final points to the initial points in such a way that no new crossings are created in the projection along the x-axis [43].

By enforcing closure in a way that yields a pure braid, we are able to apply Proposition 2.24 to extract components of a reducible braid.

Once the algebraic braid is computed, we map each generator σ_i of the algebraic braid to its corresponding Burau matrix, given in Equation (2.2), and we instantiate the Burau matrices with a fixed real-valued t such that $t \approx 1$. (Corollary 2.27 can be used to choose the value of t.) The product $\prod B(\sigma_{b_\ell})(t \approx 1)$, where $\beta = \prod_\ell \sigma_{b_\ell}$, is the Burau matrix corresponding to the braid of trajectories.

We should note that an eigenvector found in this manner may not necessarily delineate all components of a reducible braid. In fact, for a pure braid β, all vectors of the form $(0, \cdots, 0, 1, 0, \cdots, 0)$ are eigenvectors of $B(\beta)(1)$. In practice, these vectors are easy to distinguish from eigenvectors that reveal dynamical structure.

The assiduous reader may note that most theorems in this document were proven for the more general case $t \in \mathbb{C}$. These proofs also hold for $t \in \mathbb{R}$. In practice, instantiating the Burau matrices to a real-valued $t \in \mathbb{R}$ halves the number of arithmetic operations required, and no fidelity is forfeited.

2.5   Examples

We demonstrate the relevance of our contributions on two examples: the blinking vortex flow and the (modified) Duffing oscillator.

2.5.1   Blinking vortex flow

The blinking vortex flow was introduced by Aref as an idealization of stirring [44]. The flow is given by a pair of vortices separated by a finite distance, blinking on and off periodically in an alternating fashion in an incompressible, inviscid fluid. We consider a modified version of this flow in an unbounded domain (modeled on the complex plane).

The velocity field due to a single point vortex located at $x = a$ on the x-axis is given by

$$\dot{r} = 0,$$
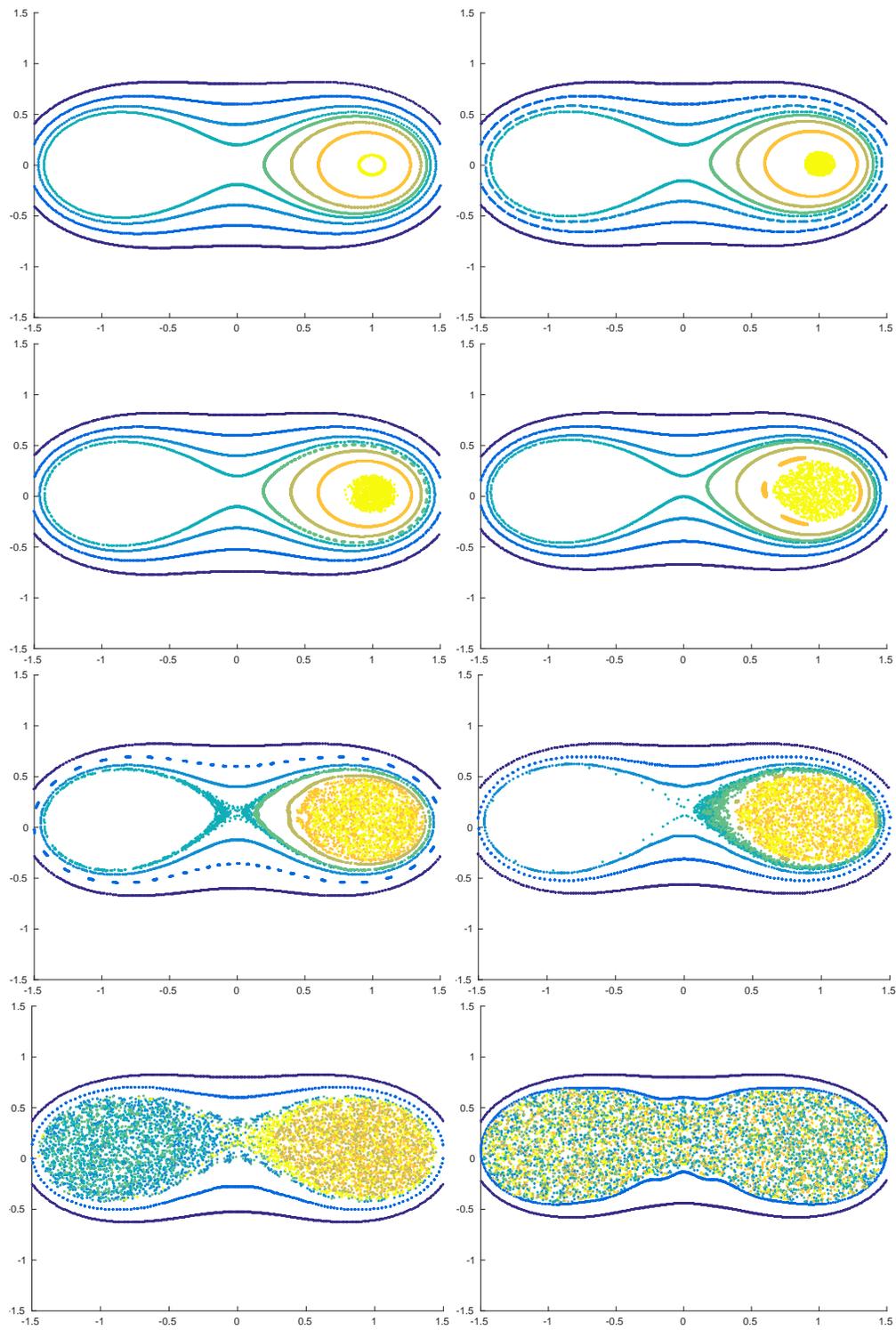$$\dot{\theta} = \frac{\Gamma}{2\pi r},$$

Figure 2.8: Illustrations of Poincaré sections for Aref's blinking vortex flow; from left to right, top to bottom: $\mu = 1, 5, 10, 20, 30, 35, 40, 50$.

where Γ is the strength of the vortex, and $r = \sqrt{(x-a)^2 + y^2}$ is the distance to the center of the vortex.

The mapping, in dimensionless form [58], induced by two identical vortices at $\xi_i = \pm a$, each acting for time T, is given by the *twist map*

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} \xi_i + (x-\xi_i)\cos\Delta\theta - y\sin\Delta\theta \\ (x-\xi_i)\sin\Delta\theta + y\cos\Delta\theta \end{pmatrix},$$

where $\Delta\theta = \frac{\mu}{r^2}$, with $\mu = \frac{\Gamma T}{2\pi a^2}$, and $r = \sqrt{(x-x_i)^2 + y^2}$. The parameter μ is the flow strength, and its value controls the behavior of the system. We make distances dimensionless with respect to a and place the vortices at $\xi_i = \pm 1$.

When both vortices act simultaneously ($T = 0$, $\mu = 0$), the system is integrable. We perturb the system by increasing μ from zero and study the Poincaré sections $t = kT, k \in \mathbb{Z}$ (See Figure 2.8). Chaotic regions appear for all $\mu > 0$ [59]. For small values of μ, small chaotic regions exist near the elliptic and hyperbolic points. As μ increases, the size of the chaotic regions grow, destroying confining KAM surfaces as the chaotic regions merge.

Using the methods described above, we give eigenvectors for the blinking vortex flow for $\mu = 0.01, 0.05, 0.10, 0.20, 0.35, 0.50$ in Figure 2.9.

2.5.2   Modified Duffing oscillator

In this example, we study a modified Duffing oscillator, given by

$$\begin{aligned} \dot{x} &= y + \alpha\cos(\omega t), \\ \dot{y} &= x\left(1 - x^2\right) - \delta y + \gamma\cos(\omega t), \end{aligned} \tag{2.5}$$

with $\alpha = 0.1$, $\gamma = 0.14$, $\delta = 0.08$, $\omega = 1$.

This compressible system is also studied by Allshouse and Thiffeault [36] with the same parameters as an example a system with two primary regions of mixing: (i) the limit cycle and its basin of attraction (yellow in Figure 2.10) and (ii) the rest of the domain (blue/green).

A *limit cycle* (Figure 2.11) is an isolated closed trajectory [60]. A trajectory is *isolated* if neighboring trajectories are not closed; they spiral toward or away from the limit cycle.
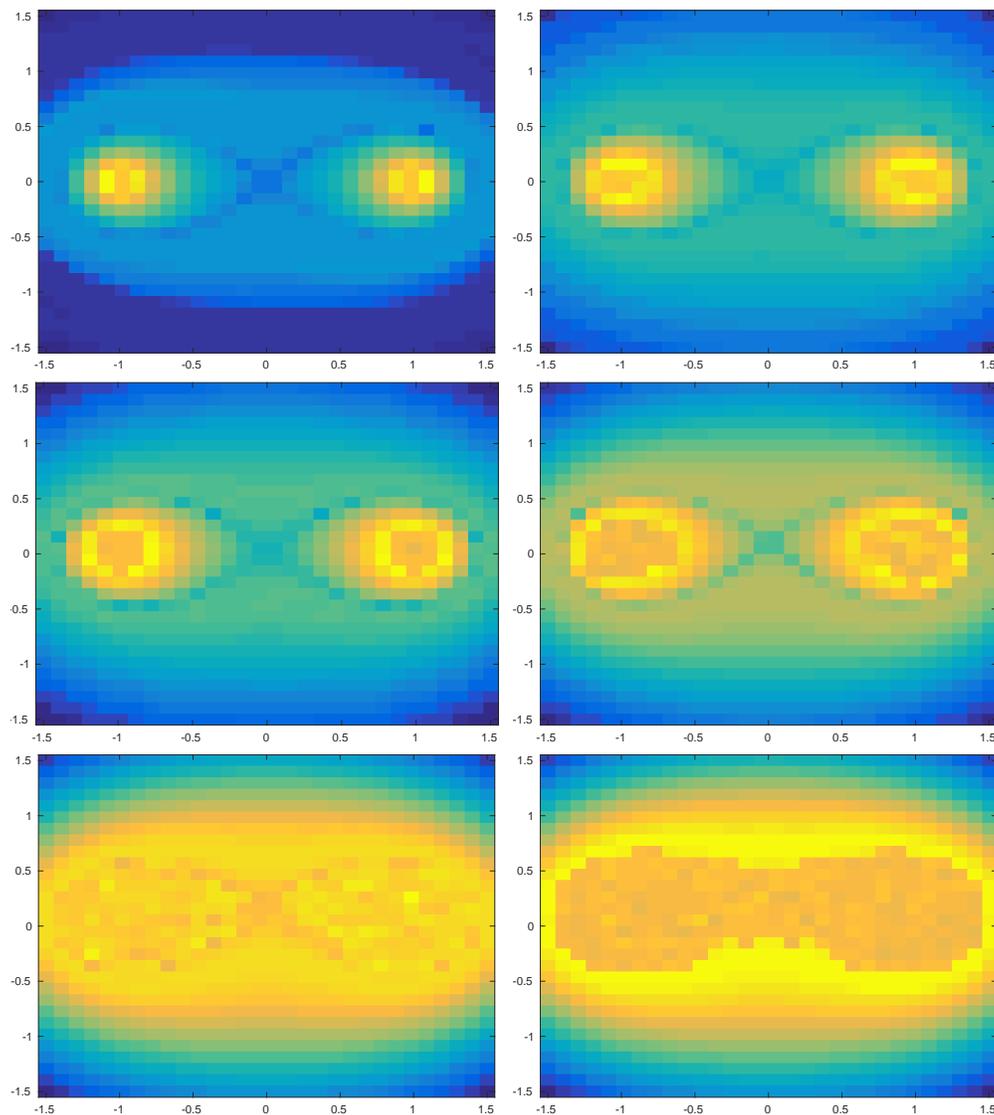
Figure 2.9: Eigenvectors of Burau matrix computed from trajectories of Aref's blinking vortex flow; from left to right, and top to bottom: $\mu = 0.01, 0.05, 0.10, 0.20, 0.35, 0.50$.

We say that a limit cycle is *stable* (or *attracting*) if all neighboring trajectories approach the limit cycle. Otherwise, we say that the limit cycle is *unstable*. A stable limit cycle is an example of an attractor.

Informally, an attractor is a set to which all neighboring trajectories converge. The concept of an attractor generated great interest when Ruelle and Takens suggested that turbulent behavior in fluids might be due to the presence of "strange" attractors [61]. Although attractors play an important role in
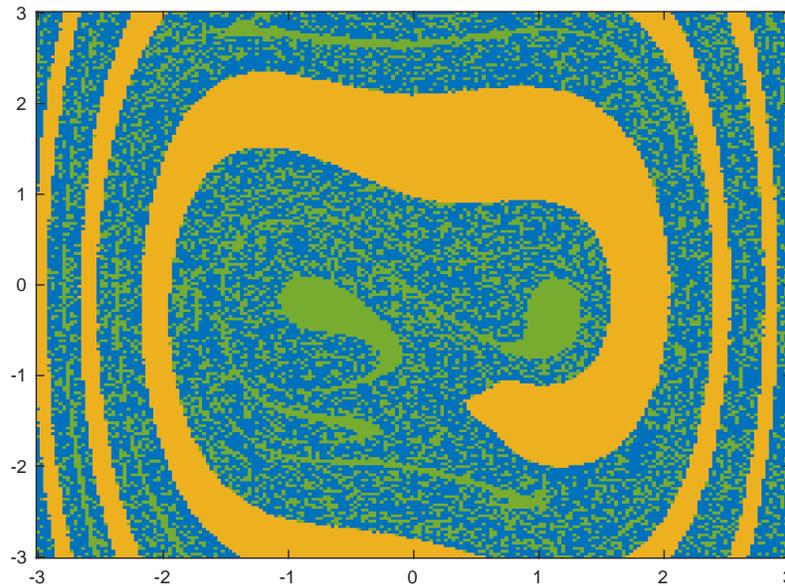
Figure 2.10: **Attractors of the modified Duffing oscillator.** A coloring of initial particle positions by their position at time $t = 250$.

dynamical systems, there is no universal agreement over the most useful definition. We refer the reader to [62] and [61] for a discussion of the literature.

Following the exposition of [60], we say that a closed set A is an *attractor* if the following are true:

(1) A is an invariant set: any trajectory $x(t)$ that starts in A stays in A for all time;

(2) A attracts an open set of initial conditions: there is an open set $U \supset A$ such that if $x(0) \in U$, then the distance from $x(t)$ to A tends to zero as $t \to \infty$. The largest such U is called the *basin of attraction* of A;

(3) A is minimal: there is no proper subset of A satisfying conditions (1) and (2).

Trajectories of the modified Duffing oscillator belong to one of three types, each illustrated in Figure 2.12 with the corresponding colors. The yellow region corresponds to the yellow limit cycle and its basin of attraction. The green regions attract to one of the two green limit cycles. The blue region
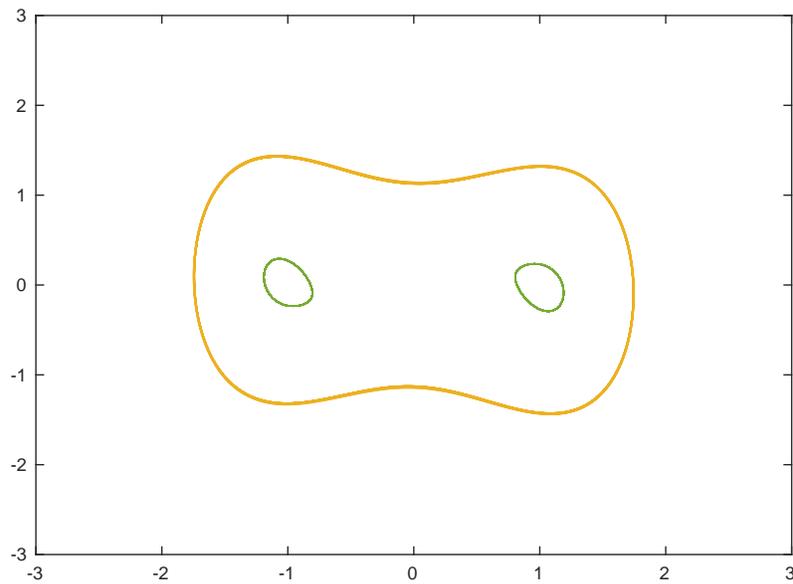
Figure 2.11: Three limit cycles of the modified Duffing oscillator.

attracts to neither the yellow limit cycle nor the green limit cycles during the time period studied.
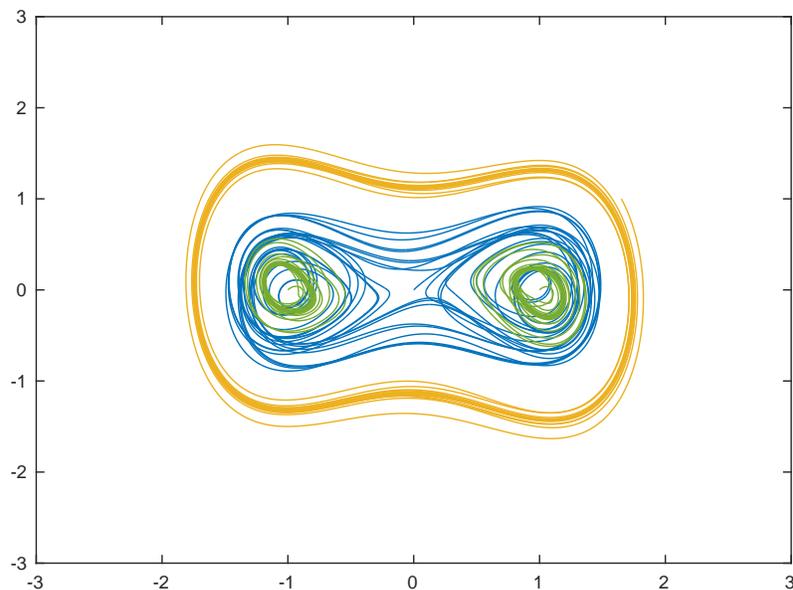

Figure 2.12: Phase portrait of the modified Duffing oscillator

Using the methods described above, we give eigenvectors for the modified Duffing oscillator in Figure 2.13. In addition to the two initial conditions found by Allshouse and Thiffeault [36], we are able to detect two additional

limit cycles (green in Figure 2.10).

We note that since the modified Duffing oscillator is a compressible system, particle positions can in fact coincide. In order to form a well-defined braid, we select a sufficiently sparse sampling of the domain and a time window so that the dynamics are reasonably well-captured yet no two particles coincide at any time. Allshouse and Thiffeault [36] argue that the ability of the braid theoretic approach to detect coherent sets even in an incompressible flow is a testament to the wide applicability of the method.
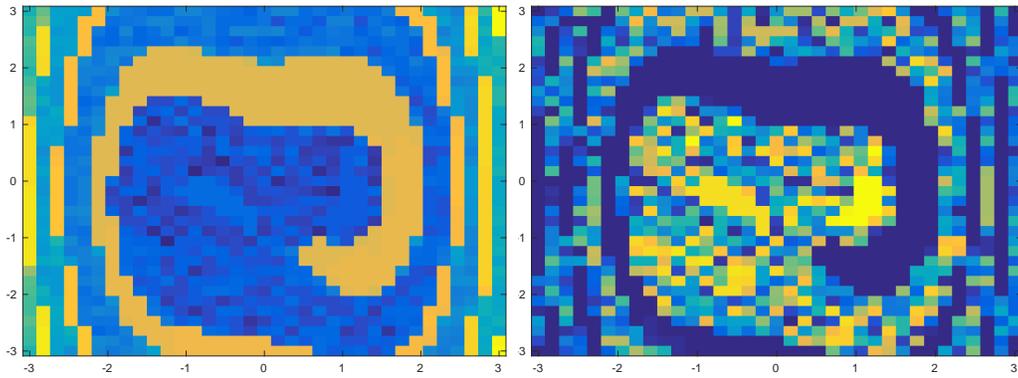

Figure 2.13: Eigenvectors of the Burau matrix computed from trajectories of the modified Duffing oscillator depict three types of initial condition.

2.6   Discussion and future directions

2.6.1   A faithful representation of the braid group

In this work, we have chosen to use the Burau representation for the analysis of spatially sparse particle trajectories. The Burau representation is not faithful but confers computational advantages in both space and time, compared to the faithful Lawrence-Krammer representation, which is given by $\frac{n(n-1)}{2} \times \frac{n(n-1)}{2}$ matrices [63]. If a greater level of topological and dynamical detail is preferred, we may wish to consider the Lawrence-Krammer representation instead.

2.6.2   Parallelism

The algorithm that we describe in this chapter lends itself naturally to parallelism. We do not discuss the details here, but we note that the matrix chain multiplication required by our algorithm is amenable to optimization. In particular, in addition to the usual considerations of the matrix chain mul-
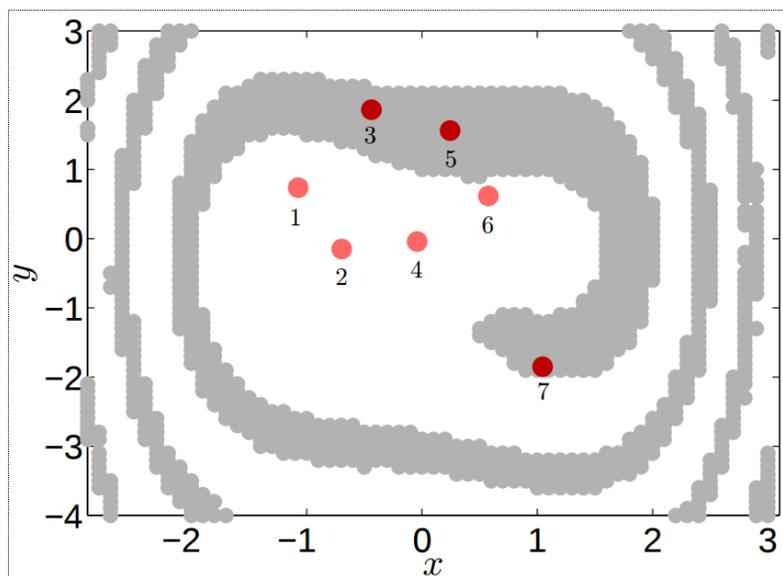
Figure 2.14: Allshouse and Thiffeault detect two types of initial conditions for the modified Duffing oscillator [36]. The dots (numbered from left to right) are the initial conditions for the trajectories that are studied further by Allshouse and Thiffeault as representative trajectories for the two types of initial conditions.

tiplication problem [64, 65], we remark that by taking the sequence of braid generators into account, then depending on the flow, we can potentially partition the braid $\beta = \prod_{\ell}^{L} \sigma_{b_\ell}$ into subsequences of neighboring generators (e.g., each subsequence consists only of generators $\sigma_{i-j}, \sigma_{i-j+1}, \cdots, \sigma_{i+j}$, for small j). This effectively partitions the sequence of matrix multiplications into subsequences that each consist of (mostly) sparse matrix multiplications, thereby reducing the computational requirements of our analysis. Future work may wish to formalize these and other computational considerations.

Chapter 3

TOPOLOGICAL DATA ANALYSIS

3.1 Introduction

The proliferation of sensors and advancement of technology has led to the production and collection of unprecedented amounts of data in recent years. The data are often noisy, non-linear, and high-dimensional, and the effectiveness of traditional tools may be limited.

Recently, there has been a lot of interest in and success with developing topologically-motivated techniques to analyze high-dimensional data. These approaches are especially useful when a topological method is sensitive to large- and small-scale features that might not be detected by analysis methods that may obscure geometric features, such as principal component analysis (PCA), multi-dimensional scaling (MDS), and cluster analysis. Topological data analysis has found success in many applications, including the identification of finer stratifications of breast cancer patients, voting patterns of the House of Representatives, and playing styles of NBA players [66]; the modeling and forecast of contagions on networks [67]; and the detection of intruders in sensor networks [68].

In many situations when we are given data points in the form of very long vectors, only a few coordinates or dimensions might be relevant to the question(s) of interest. The data may have been generated by an underlying system with far fewer degrees of freedom than the ambient dimension might suggest, and the data may live on or, in the presence of noise, near a submanifold of the ambient space. Manifold learning is concerned with estimating the geometric or topological properties of the submanifold of interest from the points given.

In this section, we are interested in computing topological invariants, namely homology, that provide a good characterization of the submanifold from the data points given. Niyogi, Smale, and Weinberger [69] proved that in the case where data are drawn by sampling a probability distribution that has

support on or near a submanifold of Euclidean space, it is possible to "learn" the homology of the submanifold with high confidence.

We begin by defining some simplicial complexes that can be used to summarize the shape of a space X. We also provide an overview of previous work on Reeb graphs and their higher-dimensional analogues, Reeb spaces, which allow us to summarize the structure of a topological space X with respect to level sets of a continuous function $f : X \to \mathbb{R}^d, d \geq 1$.

Reeb graphs and Reeb spaces can also be adapted for the setting of point clouds using a construction called mapper, which was introduced by Singh, Mémoli, and Carlsson [5]. In their preprint, Carrière and Oudot relate 1-dimensional mapper constructions of a topological space X and a continuous real-valued function $f : X \to \mathbb{R}$ to the corresponding Reeb graph and show that 1-dimensional mapper constructions are stable under perturbations. Munch and Wang use category theory to provide a more general framework and relate mapper constructions for a topological space X and a continuous multivariate function $f : X \to \mathbb{R}^d, d \geq 1$, to the corresponding Reeb space. They show that categorical Reeb spaces and categorical mapper constructions converge under an interleaving distance between their categorical representations.

3.2   Contributions

As we provide an overview of previous work, we use the techniques of de Silva et al. [20] to show that the interleaving distance between mapper constructions for continuous maps $f : X \to \mathbb{R}^d$, $g : X \to \mathbb{R}^d$, $d \geq 1$, from a topological space X to a d-dimensional parameter space is stable with respect to the supremum norm (Proposition 3.21).

We then extend Munch and Wang's work. We propose that the data of mapper constructions over point clouds can be stored as a functor. Following the advice of Singh, Mémoli, and Carlsson [5], we consider clusters to be the discrete analog of the path-components studied by Munch and Wang [19]. For each pair $f = (X, f)$, where X is a finite set and $f : X \to \mathbb{R}^d$ is a filter function into a d-dimensional parameter space, and for a given finite open cover $\mathcal{U}$ of the image of f, the nerve $K = N(\mathcal{U})$, and clustering parameter $\delta \geq 0$, we associate a functor $\dot{C}^f_{K,\delta} : \mathbf{Cell}(K)^{\mathrm{op}} \to \mathbf{Set}$, called the abstract

mapper from the category of $\textbf{Cell}(K)^{\text{op}}$.

For each fixed δ, the functor $\dot{C}^f_{K,\delta}$ corresponds to a mapper construction at a fixed clustering resolution. In order to compare mapper constructions over a range of clustering resolutions, we introduce the hierarchical abstract mapper, which is given by a family of abstract mappers $\left\{\dot{C}^f_{K,\delta}\right\}_{\delta \geq 0}$.

Now, each functor $\dot{C}^f_{K,\delta} : \textbf{Cell}(K)^{\text{op}} \to \textbf{Set}$ can be pushed into the category $\textbf{Set}^{\textbf{Open}(\mathbb{R}^d)}$ and viewed as a functor $\mathring{C}^f_{K,\delta} : \textbf{Open}(\mathbb{R}^d) \to \textbf{Set}$, which enables us to define an interleaving distance between hierarchical abstract mappers $\mathring{\mathfrak{C}}^f_K = \left\{\mathring{C}^f_{K,\delta}\right\}_{\delta \geq 0}$ over point clouds.

We prove that, as the resolution of the cover $\mathcal{U}$, given by

$$\text{res}(\mathcal{U}) := \sup\left\{\text{diam}(U_\alpha) \mid U_\alpha \in \mathcal{U}\right\},$$

goes to zero, the interleaving distance between abstract mappers over point clouds are stable in two aspects:

Theorem 3.38. *Let (X, d) be a finite metric space. Let $f, g : X \to \mathbb{R}^d$ be filter functions into a d-dimensional parameter space $\mathbb{R}^d$. Then as $\text{res}(\mathcal{U}) \to 0$, where $\mathcal{U}$ is a finite open cover for $\text{image}(f) \cup \text{image}(g)$, we have that*

$$d^{\mathcal{I}}(\mathring{C}^f_{K,\delta}, \mathring{C}^g_{K,\delta}) \leq \|f - g\|_\infty$$

for all $\delta \geq 0$.

Theorem 3.41. *Let X and Y be compact subsets of a metric space (Z, d). Let $f : X \cup Y \to \mathbb{R}^d$ be a Lipschitz continuous filter function into $\mathbb{R}^d$, with Lipschitz constant $L \in \mathbb{R}$. Denote the restrictions of f to X and Y by $f_X : X \to \mathbb{R}^d$ and $f_Y : Y \to \mathbb{R}^d$, respectively. Then as $\text{res}(\mathcal{U}) \to 0$, where $\mathcal{U}$ is an open cover for $\text{image}(f) \cup \text{image}(g)$, we have that*

$$d^{\mathcal{I}}(\mathring{\mathfrak{C}}^{f_X}_K, \mathring{\mathfrak{C}}^{f_Y}_K) \leq \max\left(d^{\mathcal{H}}(X, Y), \, L \cdot d^{\mathcal{H}}(X, Y)\right),$$

where $d^{\mathcal{H}}(X, Y)$ is the Hausdorff distance between X and Y.

Theorem 3.41 relates mapper constructions over point clouds to mapper constructions over topological spaces.

Furthermore, we show that hierarchical abstract mappers correspond to dendrograms of single-linkage hierarchical clustering when we take the filter function $f : X \to \mathbb{R}$ to be a constant function. In these cases, we show that the interleaving distance between hierarchical abstract mappers over two different sets X and Y is bounded above by the Gromov-Hausdorff distance between the dendrograms over X and Y.

Finally, we give an algorithm for hierarchical mapper constructions, which allow the study of topological features of mapper constructions over a range of clustering resolutions, enabling the analysis of topological features using statistical methods. We apply our tools to the analysis of high-dimensional geospatial sensor data and provide a statistic for quantifying climate anomalies.

3.3 Simplicial complexes

It is desirable to work with simplicial complexes because the homology of a simplicial complex is easily computable [70, 71].

Here, we give a brief description of simplicial complexes. We refer the reader to [72, 73] for a more in-depth discussion.

Abstract simplicial complexes

Definition 3.1 (abstract simplicial complex)**.** An *abstract simplicial complex* is a collection Σ of finite non-empty sets, such that if $\sigma \in \Sigma$, and τ is a non-empty subset of σ, then $\tau \in \Sigma$.

An element σ of Σ is a *simplex* of Σ. Each non-empty subset τ of a simplex σ is a *face of σ*. The *vertex set V* of Σ is the collection of the one-point elements of Σ. We do not distinguish between the vertex $v \in V$ and the 0-simplex $\{v\} \in \Sigma$.

The *dimension of a simplex* is one less than the number of its elements. The *dimension of a simplicial complex* is the greatest dimension of its simplices or is infinite if there is no such greatest dimension.

A subcollection of Σ that is itself a simplicial complex is called a *subcomplex* of Σ.

Nerve and Nerve Lemma

Given a topological space X, there are many different simplicial complexes that can be constructed from it. One particularly useful construction that we will devote our attention to is the *nerve*.

Definition 3.2 (nerve)**.** Let $\mathcal{A}$ be a collection of subsets of the space X. The *nerve* of $\mathcal{A}$, denoted $N(\mathcal{A})$, is the abstract simplicial complex with vertices consisting of the elements of $\mathcal{A}$ and simplices given by the finite subcollections $\{A_{i_0}, \cdots, A_{i_n}\}$ of $\mathcal{A}$ such that

$$A_{i_0} \cap \cdots \cap A_{i_n} \neq \emptyset.$$

When $\mathcal{A}$ is a covering of X, the nerve $N(\mathcal{A})$ can be thought of as a combinatorial representation of X.

The nerve of X is a useful construction because the Nerve Lemma [74, 75] provides criteria that guarantee that the nerve of a covering of X is homotopy equivalent to X.

Theorem 3.3 (Nerve Lemma [73])**.** *Let X be a paracompact space. If $\mathcal{U}$ is an open cover of X such that every non-empty intersection of finitely many sets in $\mathcal{U}$ is contractible, then X is homotopy-equivalent to the nerve $N(\mathcal{U})$.*

Throughout this thesis, we assume all covers are good open covers so that the Nerve Lemma applies.

Čech complexes

When X is a metric space, we can consider the covering of X by ε-balls to form a *Čech complex*.

Definition 3.4 (Čech complex)**.** Let (X, d) be a metric space. For $\varepsilon > 0$, let $A \subseteq X$ such that $X = \bigcup_{a \in A} B_\varepsilon(a)$, where $B_\varepsilon(a) := \{x \in X : d(x, a) < \varepsilon\}$ is a ball of radius ε centered at $a \in A$. The *Čech complex* of (A, ε), denoted $\check{C}(A, \varepsilon)$, is the nerve of the covering $\{B_\varepsilon(a)\}_{a \in A}$.

The Čech complex can be used to approximate the homology of a manifold.

Theorem 3.5 ([76])**.** *Let M be a compact Riemannian manifold. There exists a positive number e so that $\check{C}(M, \varepsilon)$ is homotopy equivalent to M whenever $\varepsilon \leq e$.*

Moreover, for every $\varepsilon \leq e$, there is a finite subset $V \subseteq M$ so that the subcomplex $\check{C}(V, \varepsilon) \subseteq \check{C}(M, \varepsilon)$ is also homotopy equivalent to M.

Vietoris-Rips complexes

Unfortunately, the Čech complex can be computationally intensive to compute. One can consider the Vietoris-Rips complex as an alternative, which allows us to consider only pairs instead of all subcollections.

Definition 3.6 (Vietoris-Rips complex)**.** Let (X, d) be a metric space. The Vietoris-Rips complex $VR(X, \varepsilon)$ is the simplicial complex with vertex set X, such that a set $\{x_0, \cdots, x_k\}$ spans a k-simplex if and only if $d(x_i, x_j) \leq \varepsilon$ for all $0 \leq i, j \leq k$.

We can relate the Čech complex and Vietoris-Rips complex through the following inclusions:

Proposition 3.7 ([76])**.** *For $\varepsilon > 0$, we have the inclusions*

$$\check{C}(X, \varepsilon) \subseteq VR(X, \varepsilon) \subseteq \check{C}(X, 2\varepsilon).$$

3.4   Persistent homology

Persistent homology is motivated by the desire to distinguish noise from features in data and the need to examine data over a range of scales. We use persistent homology to measure the scale, or resolution, of a topological feature.

We follow the expositions of Edelsbrunner and Harer [77, 78] as we introduce background material for persistent homology. Although persistence can be defined for any sequence of vector spaces connected by homomorphisms, in our applications, we are only concerned with persistent homology for simplicial complexes. We refer the reader to [72, 73] for a thorough discussion of homology.

Throughout this chapter, we consider only homology with $\mathbb{Z}/2\mathbb{Z}$ coefficients. We write $H_r(X)$ for $H_r(X; \mathbb{Z}/2\mathbb{Z})$, and we call its rank the *rth Betti number of X*, denoted $\beta_r(X)$.

Filtrations

In order to define persistent homology for simplicial complexes, we introduce the notion of a *filtration* of a simplicial complex given by a *monotonic* map.

Definition 3.8 (monotonic)**.** Let K be a simplicial complex. We say that a function $f : K \to \mathbb{R}$ is *monotonic* if it is non-decreasing along increasing chains of faces; that is, $f(\tau) \leq f(\sigma)$ whenever τ is a face of σ.

If $f : K \to \mathbb{R}$ is monotonic, then the sublevel set

$$K(a) = f^{-1}(-\infty, a]$$

is a subcomplex of K for every $a \in \mathbb{R}$.

In particular, if K is a simplicial complex with m simplices, then we can find an increasing sequence of $n + 1 \leq m + 1$ different subcomplexes:

$$\emptyset = K_0 \subseteq K_1 \subseteq \cdots \subseteq K_n = K.$$

Letting $a_1 < \cdots < a_n$ be the images of the simplices in K under $f : K \to \mathbb{R}$, and $a_0 = -\infty$, we have $K_i = K(a_i)$ for each i.

We call this nested sequence of subcomplexes a *filtration* of (K, f).

The inclusion maps $K_{i-1} \to K_i$ between subcomplexes induce the following homomorphisms by functorality:

$$0 = H_*(K_0) \to H_*(K_1) \to \cdots \to H_*(K_m) = H_*(K).$$

As simplices get added to K_{i-1} to form K_i, new homology classes may come into existence, and existing homology classes may become trivial or merge with one another.

Persistent homology groups

Definition 3.9 (*p*th persistent homology groups)**.** Let $f_p^{i,j} : H_p(K_i) \to H_p\left(K_j\right)$ be the homomorphism induced by the inclusion $K_i \to K_j$, for $i \leq j$. The *pth persistent homology groups* are the images

$$H_p^{i,j} = \mathrm{im}\, f_p^{i,j},$$

for $0 \leq i \leq j \leq n$.

The ranks of these groups are called the *pth persistent Betti numbers*

$$\beta_p^{i,j} = \operatorname{rank} H_p^{i,j}.$$

Note that the persistent homology groups $H_p^{i,j}$ consist of homology classes of K_i that still exist in K_j, and $H_p^{i,i} = H_p\left(K_i\right)$.

A homology class $\alpha \in H_p\left(K_i\right)$ is *born* at K_i if it is not in the image of the map induced by the inclusion $K_{i-1} \subset K_i$. We say α *dies entering* K_j if the image of the map induced by $K_{i-1} \subset K_{j-1}$ does not contain the image of α but the image of the map induced by $K_{i-1} \subset K_j$ does. The *persistence* of α is $a_j - a_i$ or $j - i$, depending on the application. A class is *essential* if it does not die within the filtration.

Persistence diagrams

A persistence diagram is a multiset of points in the extended plane $\mathbb{R}_\infty^2$.

Let $\mu_p^{i,j}$ be the number of p-dimensional classes born at K_i and dying entering K_j. Then the *pth persistence diagram* of the filtration of (K, f), denoted $\operatorname{dgm}_p(f)$, consists of points (a_i, a_j), each with multiplicity $\mu_p^{i,j}$.

3.5   Reeb graphs and Reeb spaces

3.5.1   Reeb graphs

In addition to studying the shape of a topological space X, we can also study properties of a topological space X through the level sets of a continuous function $f : X \to \mathbb{R}$. A tool for tracking and visualizing the path-components of the level sets of f is the *Reeb graph*. Since Reeb graphs depend on the choice of map $f : X \to \mathbb{R}$, different functions can be used to study different properties of a space.

Reeb graphs were originally introduced in the context of Morse theory [79]. They can be used to select meaningful level sets and have been used as a tool for shape comparison [80, 81], data skeletonization [82], and surface denoising [83]. Furthermore, since Reeb graphs are fast to compute [84, 85], they can be useful for analyzing very large datasets.

Geometric Reeb graphs

The geometric Reeb graph provides an intuitive summary of the structure of the level sets of a pair (X, f).

Definition 3.10 (geometric Reeb graph [20])**.** Let X be a topological space, and let $f : X \to \mathbb{R}$ be a continuous real-valued function. For $x, x' \in X$, we write $x \sim_f x'$ if x and x' belong to the same path-component of a level set $f^{-1}(a)$, for some $a \in \mathbb{R}$. The quotient space $X/ \sim_f$ is the *(geometric) Reeb graph* of (X, f).

Example 3.11 (height function)**.** We illustrate the Reeb graph of a surface given by the height function in Figure 3.1.
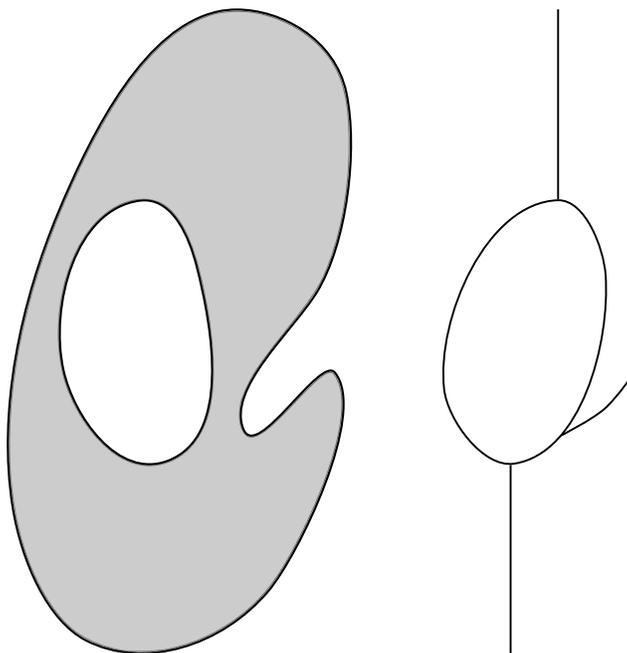
Figure 3.1: Reeb graph of height function.

In general, the quotient space $X \to X/ \sim_f$ may be poorly behaved. Thus, we restrict ourselves to a class of *constructible* pairs (X, f).

Definition 3.12 ($\mathbb{R}$-**Top** [20])**.** Let $\mathbb{R}$-**Top** be the category whose objects consist of pairs (X, f), where X is a topological space and $f : X \to \mathbb{R}$ is a continuous map, and whose morphisms $\phi : (X, f) \to (Y, g)$ are continuous

maps $\phi : X \to Y$ such that the following diagram commutes:

$$
\begin{array}{ccc}
X & \xrightarrow{\phi} & Y \\
 & {}_{f}\searrow \quad \swarrow{}_{g} & \\
 & \mathbb{R} &
\end{array}
$$

Definition 3.13 ($\mathbb{R}$-**Top**^{c} [20])**.** We say that an object of $\mathbb{R}$-**Top** is *constructible* if it is isomorphic to some (X, f) constructed in the following manner: given a finite set $A = \{a_0, \cdots, a_n\}$, listed in increasing order,

- specify a locally path-connected space V_i for each $0 \le i \le n$,

- specify a locally path-connected space E_i, for each $0 \le i < n$, and

- specify continuous maps $l_i : E_i \to V_i$ and $r_i : E_i \to V_{i+1}$.

Let X be the quotient space

$$
\left(\left(\bigsqcup_i V_i \times \{a_i\} \right) \bigsqcup \left(\bigsqcup_i E_i \times [a_i, a_{i+1}] \right) \right) / \sim,
$$

with $(l_i(x), a_i) \sim (x, a_i)$ and $(r_i(x), a_{i+1}) \sim (x, a_{i+1})$ for all i and all $x \in E_i$, and let $f : X \to \mathbb{R}$ be the projection onto the second factor.

We denote the full subcategory of constructible $(X, f) \in \mathbb{R}$-**Top** by $\mathbb{R}$-**Top**^{c}.

Examples of constructible (X, f) include Morse functions on compact manifolds and piecewise linear functions on compact polyhedra.

Abstract Reeb graphs

In many applications, it is important to ensure that two Reeb graphs are similar when the functions they arise from are similar. To do so, de Silva et al. [20] show that the data of a Reeb graph can be stored abstractly as a functor. In this setting, they define the *interleaving distance* between pairs of Reeb graphs and show that the interleaving distance is stable under perturbations. We summarize some of their definitions and results here and call on them later.

Definition 3.14 (abstract Reeb graph [20]). Let X be a topological space, and let $f : X \to R$ be a continuous real-valued function. Let $\mathbf{Open}(\mathbb{R})$ denote the category with objects consisting of open sets in $\mathbb{R}$ and arrows $I \to J$ between two objects if and only if $I \subseteq J$.

The *abstract Reeb graph* of (X, f) is a functor $F : \mathbf{Open}(\mathbb{R}) \to \mathbf{Set}$ that maps each open set $I \subseteq \mathbb{R}$ to the set of path-components of $f^{-1}(I)$, denoted $F(I)$, and each arrow $I \subseteq J$ to the set map $F(I) \to F(J)$ induced by the inclusion $f^{-1}(I) \subseteq f^{-1}(J)$, denoted $F[I \subseteq J]$.

Note that an abstract Reeb graph is an object in the category of functors $\mathbf{Set}^{\mathbf{Open}(\mathbb{R})}$.

Let $\mathcal{C} : \mathbb{R}\textbf{-Top} \to \mathbf{Set}^{\mathbf{Open}(\mathbb{R})}$ denote the functor that maps $f = (X, f)$ to its Reeb graph $\mathcal{C}(f) = F$, with

$$F(I) = \pi_0 f^{-1}(I), \quad F[I \subseteq J] = \pi_0[f^{-1}(I) \subseteq f^{-1}(J)],$$

where we let π_0 denote the set of path-components of a space.

Interleaving distance between Reeb graphs

Interleavings are approximate isomorphisms. An *isomorphism* between functors $F, G : \mathbf{Open}(\mathbb{R}) \to \mathbf{Set}$ is a pair of families of maps

$$\phi_I : F(I) \to G(I), \quad \psi_I : G(I) \to F(I)$$

that are natural with respect to inclusions $I \subseteq J$ such that ϕ_I and ψ_I are inverses for all I.

An ε-interleaving gives some leeway by allowing the codomains of ϕ_I and ψ_I to be given by an ε-expansion of I.

Definition 3.15 (ε-interleaving between Reeb graphs). For an open interval $I = (a, b) \subseteq \mathbb{R}$, let $I^\varepsilon = (a - \varepsilon, b + \varepsilon)$.

An *ε-interleaving* between two Reeb graphs $F, G : \mathbf{Open}(\mathbb{R}) \to \mathbf{Set}$ is a pair of families of maps

$$\phi_I : F(I) \to G(I^\varepsilon), \quad \psi_I : G(I) \to F(I^\varepsilon)$$

that are natural with respect to inclusions $I \subseteq J$ and satisfy

$$\psi_{I^\varepsilon} \circ \phi_I = F[I \subseteq I^{2\varepsilon}], \quad \phi_{I^\varepsilon} \circ \psi_I = G[I \subseteq I^{2\varepsilon}]$$

for all I.

$$
\begin{array}{ccc}
F(I) & \xrightarrow{\phi_I} & \\
\downarrow{\scriptstyle F[I \subseteq I^{2\varepsilon}]} & & G(I^\varepsilon) \\
& \swarrow{\scriptstyle \psi_{I^\varepsilon}} & \\
F(I^{2\varepsilon}) & &
\end{array}
\qquad
\begin{array}{ccc}
G(I) & \xrightarrow{\psi_I} & \\
\downarrow{\scriptstyle G[I \subseteq I^{2\varepsilon}]} & & F(I^\varepsilon) \\
& \swarrow{\scriptstyle \phi_{I^\varepsilon}} & \\
G(I^{2\varepsilon}) & &
\end{array}
$$

When there exists an ε-interleaving between Reeb graphs F and G, we say that F and G are *ε-interleaved*.

Definition 3.16 (interleaving distance)**.** The *interleaving distance* between two Reeb graphs $F, G : \mathbf{Open}(\mathbb{R}) \to \mathbf{Set}$ is given by

$$d^{\mathcal{I}}(F, G) = \inf \{\varepsilon \mid F, G \text{ are } \varepsilon\text{-interleaved}\}.$$

(We take the infimum of an empty set to be ∞.)

De Silva et al. [20] show that the interleaving distance on the space of Reeb graphs is an extended pseudometric. (It takes values in $[0, \infty]$, with $d^{\mathcal{I}}(\mathcal{C}(f), \mathcal{C}(f)) = 0$, is symmetric, and satisfies the triangle inequality.)

3.5.2   Reeb spaces

We can also study higher-dimensional analogues of Reeb graphs, called *Reeb spaces*. Reeb spaces are generalizations of Reeb graphs, obtained when a real-valued map $f : X \to \mathbb{R}$ is replaced by a multivariate map $f : X \to \mathbb{R}^d$.

Munch and Wang define Reeb spaces using categorical tools and show that categorical representations of Reeb spaces and categorical representations of mapper constructions converge in the interleaving distance [19]. We summarize their definitions and some of their results here and reference them in later sections.

Geometric Reeb spaces

Definition 3.17 (geometric Reeb space [19])**.** Let X be a topological space, and let $f : X \to \mathbb{R}^d$ be a continuous function. For $x, x' \in X$, we write $x \sim_f x'$ if x and x' belong to the same path-component of a level set $f^{-1}(a)$, for some $a \in \mathbb{R}^d$. The quotient space $X/\sim_f$ is the *geometric Reeb space* of (X, f).

Abstract Reeb spaces

Definition 3.18 (abstract Reeb space [19]). Let X be a topological space, and let $f : X \to \mathbb{R}^d$ be a continuous function. Let **Open**$(\mathbb{R}^d)$ denote the category with objects consisting of open sets in $\mathbb{R}^d$ and arrows $I \to J$ between two objects if and only if $I \subseteq J$.

The *abstract Reeb space* of (X, f) is a functor $F : \mathbf{Open}(\mathbb{R}^d) \to \mathbf{Set}$ that maps each open set $I \subseteq \mathbb{R}^d$ to the set of path-components of $f^{-1}(I)$, denoted $F(I)$, and each arrow $I \subseteq J$ to the set map $F(I) \to F(J)$ induced by the inclusion $f^{-1}(I) \subseteq f^{-1}(J)$, denoted $F[I \subseteq J]$.

Note that an abstract Reeb space is an object in the category of functors $\mathbf{Set}^{\mathbf{Open}(\mathbb{R}^d)}$.

Let $\mathbb{R}^d$-**Top** be the category whose objects consist of pairs (X, f), where X is a topological space and $f : X \to \mathbb{R}^d$ is a continuous map, and whose arrows are function-preserving maps $(X, f) \to (Y, g)$.

As before, we let $\mathcal{C} : \mathbb{R}^d\text{-}\mathbf{Top} \to \mathbf{Set}^{\mathbf{Open}(\mathbb{R}^d)}$ denote the functor that maps $f = (X, f)$ to its Reeb space $\mathcal{C}(f) = F$, with

$$F(I) = \pi_0 f^{-1}(I), \quad F[I \subseteq J] = \pi_0[f^{-1}(I) \subseteq f^{-1}(J)].$$

Interleaving distance between Reeb spaces

Definition 3.19 (ε-interleaving between Reeb spaces [19]). For an open set $I \subseteq \mathbb{R}^d$, let $I^\varepsilon = \{x \in \mathbb{R}^d : d(x, I) < \varepsilon\}$ denote the ε-expansion of I. An ε-*interleaving* between two abstract Reeb spaces $F, G : \mathbf{Open}(\mathbb{R}^d) \to \mathbf{Set}$ is a pair of families of maps

$$\phi_I : F(I) \to G(I^\varepsilon), \quad \psi_I : G(I) \to F(I^\varepsilon)$$

that are natural with respect to inclusions $I \subseteq J$ and satisfy

$$\psi_{I^\varepsilon} \circ \phi_I = F[I \subseteq I^{2\varepsilon}], \quad \phi_{I^\varepsilon} \circ \psi_I = G[I \subseteq I^{2\varepsilon}]$$

for all I.

$$
\begin{array}{ccc}
F(I) & \xrightarrow{\ \phi_I\ } & \\
\Big\downarrow{\scriptstyle F[I\subseteq I^{2\varepsilon}]} & & G(I^\varepsilon) \\
 & \swarrow{\scriptstyle \psi_{I^\varepsilon}} & \\
F(I^{2\varepsilon}) & &
\end{array}
\qquad
\begin{array}{ccc}
G(I) & \xrightarrow{\ \psi_I\ } & \\
\Big\downarrow{\scriptstyle G[I\subseteq I^{2\varepsilon}]} & & F(I^\varepsilon) \\
 & \swarrow{\scriptstyle \phi_{I^\varepsilon}} & \\
G(I^{2\varepsilon}) & &
\end{array}
$$
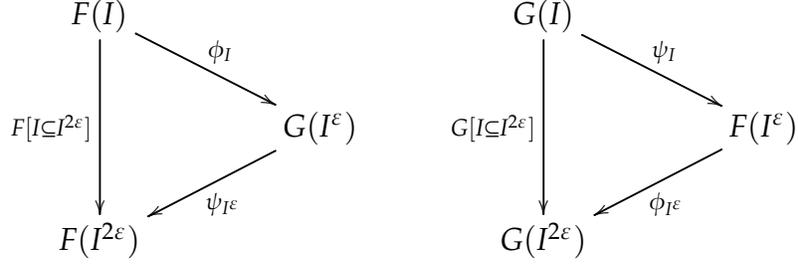
Where there exists an ε-interleaving between Reeb spaces F and G, we say that F and G are *ε-interleaved*.

Definition 3.20 (interleaving distance [19])**.** The *interleaving distance* between two Reeb spaces $F, G : \mathbf{Open}(\mathbb{R}^d) \to \mathbf{Set}$ is given by

$$
d^{\mathcal{J}}(F, G) = \inf \{\varepsilon \mid F, G \text{ are } \varepsilon\text{-interleaved}\} .
$$

(We take the infimum of an empty set to be ∞.)

Munch and Wang [19] show that the interleaving distance on the space of abstract Reeb spaces is an extended pseudometric.

Stability of interleaving distance for Reeb spaces

De Silva et al. [20] prove a stability result for interleaving distance on the space of Reeb graphs. At the time of writing, an analogous result for Reeb spaces has not been recorded. The ideas used to prove stability for Reeb graphs carry over to the setting of Reeb spaces. We state the analogous result for Reeb spaces and give a proof here.

Proposition 3.21. *Let $f = (X, f), g = (Y, g) \in \mathbb{R}^d\text{-}\mathbf{Top}$. Then*

$$
d^{\mathcal{J}}(\mathcal{C}(f), \mathcal{C}(g)) \leq \|f - g\|_\infty.
$$

Proof.

Suppose $\|f - g\|_\infty \leq \varepsilon$.

Let $I \subseteq \mathbb{R}^d$ be an open set. Since $\|f - g\|_\infty \leq \varepsilon$, we have that $\|f(x) - g(x)\| \leq \varepsilon$ for all $x \in X$. In particular, if $x \in f^{-1}(I)$, then $f(x) \in I$ and $g(x) \in I^\varepsilon$. Thus,

$$
f^{-1}(I) \subseteq g^{-1}(I^\varepsilon).
$$

Similarly, we have

$$g^{-1}(I) \subseteq f^{-1}(I^{\varepsilon}).$$

Thus, we can define

$$\phi_I = \pi_0[f^{-1}(I) \subseteq g^{-1}(I^{\varepsilon})]$$

and

$$\psi_I = \pi_0[g^{-1}(I) \subseteq f^{-1}(I^{\varepsilon})].$$

Naturality with respect to inclusions $I \subseteq J$ follows immediately.

Furthermore, since inclusions commute, we have that

$$\psi_{I^{\varepsilon}} \circ \phi_I = F[I \subseteq I^{2\varepsilon}], \quad \phi_{I^{\varepsilon}} \circ \psi_I = G[I \subseteq I^{2\varepsilon}]$$

for all I.

Thus, there is an ε-interleaving between $\mathcal{C}(f)$ and $\mathcal{C}(g)$ for all $\varepsilon \geq \|f - g\|_{\infty}$. It follows that

$$d^{\mathcal{I}}(\mathcal{C}(f), \mathcal{C}(g)) \leq \|f - g\|_{\infty}.$$

$\square$

3.6 Mapper constructions

The mapper algorithm was introduced by Singh, Mémoli, and Carlsson [5] to produce a construction that summarizes the topological properties of a point cloud.

Carriére and Oudot [76] describe 1-dimensional mapper constructions as a pixelized version of Reeb graphs, appropriate for studying point clouds.

Singh, Mémoli, and Carlsson [5] state without proof that in the case where $f : X \to \mathbb{R}$ is a continuous real-valued function, the mapper construction is a stochastic version of the Reeb graph. For a sufficiently fine covering of image(f), the mapper construction recovers the Reeb graph of f precisely.

In the more general setting where $f : X \to \mathbb{R}^d$ is a continuous multivariate function from a topological space X, Munch and Wang [19] give a formal convergence between Reeb spaces and mapper constructions in the interleaving distance.

In this section, we provide an overview of mapper constructions as they were originally introduced by Singh, Mémoli, and Carlsson [5].

3.6.1 Classical mapper construction for topological spaces

The mapper construction was introduced as a simplicial complex, given by the nerve of the pullback of an open cover by a continuous function [5].

Definition 3.22 (classical mapper construction for topological spaces [76]). Let X be a topological space. Given a continuous map $f : X \to Y$ and a finite open cover $\mathcal{U}$ of Y, the *mapper construction over X* is the nerve of the pullback $f^*(\mathcal{U})$, where we write $f^*(\mathcal{U})$ for the covering of X given by the collection of path-components of $f^{-1}(U)$, $U \in \mathcal{U}$.

3.6.2 Classical mapper construction for point clouds

In the setting of point-clouds, we consider clustering as the discrete counterpart to path-components of a topological space [76]. In this document, we use single-linkage hierarchical clustering as our clustering algorithm of choice, but depending on the application, other clustering algorithms may be appropriate as well.

Definition 3.23 (single-linkage clustering [21, 86]). Let (X, d_X) be a finite metric space. For each $\delta \geq 0$, define a relation $\sim_\delta$, where $x \sim_\delta x'$ if and only if there is a sequence $x_0, \cdots, x_k \in X$ so that $x_0 = x$, $x_k = x'$, and $d_X(x_i, x_{i+1}) \leq \delta$ for all i.

Note that $\sim_\delta$ is in fact an equivalence relation.

Furthermore, the set of clusters given by $\sim_\delta$ is exactly the set of path-components of the Vietoris-Rips complex $VR(X, \delta)$ [76], which we use to define the mapper construction for point clouds.

Definition 3.24 (classical mapper construction for point clouds [76]). Given a point cloud X, a map $f : X \to Y$, and a finite open cover $\mathcal{U}$ of Y, let $\mathcal{V}$ denote the set of path-components of the Vietoris-Rips complexes $V(f^{-1}(U), \delta)$, $U \in \mathcal{U}$. The *mapper construction over X* is the nerve of $\mathcal{V}$.

The map $f : X \to Y$ is often referred to as a *filter function* for X [5, 66] or a *reference map* to the *reference space* Y [76]. Oftentimes, f is a real-valued function that captures geometric or statistical properties of the data. We can also choose f to be application-specific [6].

Examples of filter functions. The choice of filter function(s) is critical to the successful extraction of features of relevance. A different filter function may result in a simplicial complex with a different shape, highlighting different features of the data. Some filter functions may not produce any features of interest. We give a few examples of effective filter functions and highlight some corresponding applications.

- *Density estimation* using a Gaussian kernel,

$$f_r(x) = C_r \sum_y \exp\left(\frac{-d(x,y)^2}{r}\right),$$

for $r > 0$, was employed as a filter function to differentiate between type 1 and type 2 diabetes based on six quantities: age, relative weight, fasting plasma glucose, area under the plasma glucose curve for the three-hour oral glucose tolerance test (OGTT), area under the plasma insulin curve for the OGTT, and steady-state plasma glucose response [5].

- *L-infinity centrality*, given by the maximum distance to any other data point in the data set,

$$f(x) = \max_{y \in X} d(x,y),$$

was used to robustly identify subpopulations of breast cancer patients that are consistent between two different data sets of gene expression profiles [66]. The identification of subpopulations of cancer patients who might benefit from targeted therapy is the focus of much research, as there is often much variability in patient outcomes under current treatment protocols. However, it is often difficult to identify the same subpopulation in different data sets, when the data are generated on different platforms, different sets of patients, and at different times. Topological data analysis is able to overcome some of the challenges introduced by the noise and complexity of genomic data.

- The *pth eccentricity* functions,

$$E_p(x) = \left(\frac{\sum_{y \in X} d(x,y)^p}{N}\right)^{\frac{1}{p}},$$

where $1 \leq p < \infty$, and

$$E_\infty(x) = \max_{y \in X} d(x, y),$$

were initially introduced by Hamza and Krim [87] for object represen-
tation and recognition. Singh et al. [5] use $E_1(x)$ as a filter function in
order to summarize and compare shapes.

In addition to a choice of filter function $f : X \to Y$, the mapper algorithm
also requires a choice of a finite overlapping cover $\mathcal{U}$ of the image of f,
which can vary by the size of the sets in the cover and the amount of overlap
between sets.

In the setting of point clouds, the mapper construction requires an additional
choice of resolution parameter: the clustering threshold.

Singh et al. [5] propose that in order to find the number of clusters, one can
use the edge length at which each cluster is merged; the distance within
each cluster should be smaller than the distance between clusters. For each
preimage $f^{-1}(U)$, which we call a *bin*, we can select an integer k and build a
k-interval histogram of the values for each transition in the clustering. The
clustering threshold for the mapper construction is then chosen as the last
threshold before the first gap in the histogram [66]. Although this heuristic
has worked well for many data sets, it has some limitations. If the densities
of the clusters vary, only clusters with comparatively high densities will be
selected. Furthermore, it is possible to construct examples where clusters
are distributed in such a way such that the mapper algorithm will recover
the incorrect clustering [5].

In order to address these limitations, we propose *hierarchical mappers*, which
we describe abstractly as a collection of abstract mappers in Section 3.7.3
and combinatorially as a simplicial complex in Section 3.8.2.

3.7   Abstract mapper and hierarchical abstract mapper

Abstract mappers are introduced by Munch and Wang [19] to show that for
a finite open cover $\mathcal{U} = \{U_\alpha\}_{\alpha \in A}$ of the image of f, as the *resolution* of the
cover,

$$\operatorname{res}(\mathcal{U}) := \sup \{\operatorname{diam}(U_\alpha) \mid U_\alpha \in \mathcal{U}\}, \tag{3.1}$$

goes to zero, mapper constructions converge to Reeb spaces. In order to relate mapper constructions to Reeb spaces, Munch and Wang store the data of classical mapper constructions for topological spaces as a functor, which they call the *categorical mapper*. We use the terms abstract mapper and categorical mapper interchangeably in this document. We summarize Munch & Wang's definition of abstract mapper for topological spaces and their convergence results below. In our work, we introduce abstract mapper for point clouds and hierarchical abstract mapper for point clouds, in order to offer tools for studying mapper constructions for point clouds over a range of resolutions.

3.7.1   Abstract mapper for topological spaces

For a pair (X, f), where X is a topological space and $f : X \to \mathbb{R}^d$ is a continuous map, and finite open cover $\mathcal{U} = \{U_\alpha\}_{\alpha \in A}$ of the image of f, let K denote the simplicial complex for the nerve of $\mathcal{U}$.

For each simplex $\sigma \in K$, let

$$\mathcal{U}_\sigma = \bigcap_{\alpha \in \sigma} U_\alpha \tag{3.2}$$

denote the open set in $\mathbb{R}^d$ associated to σ. Note that for $\sigma \leq \tau$, we have

$$\mathcal{U}_\sigma = \bigcap_{\alpha \in \sigma} U_\alpha \supseteq \bigcap_{\beta \in \tau} U_\beta = \mathcal{U}_\tau.$$

Thus, for a pair $\sigma \leq \tau$, we are interested in the map

$$\pi_0 f^{-1}(\mathcal{U}_\sigma) \leftarrow \pi_0 f^{-1}(\mathcal{U}_\tau).$$

Let $\mathbf{Cell}(K)$ be the category whose objects consist of simplices of K and whose morphisms $\sigma \to \tau$ are given by the face relation $\sigma \leq \tau$. We consider the opposite category, $\mathbf{Cell}(K)^{\mathrm{op}}$, whose objects consist of the simplices of K and whose morphisms $\tau \to \sigma$ are given by the face relation $\sigma \leq \tau$.

Definition 3.25 (abstract mapper [19]). The *abstract mapper* of (X, f) for the finite open cover $\mathcal{U}$ of image(f) is a functor $C_K^f : \mathbf{Cell}(K)^{\mathrm{op}} \to \mathbf{Set}$ given by

$$C_K^f(\sigma) := \pi_0 f^{-1}(\mathcal{U}_\sigma).$$

Note that the abstract mapper of (X, f) is an object of the category $\mathbf{Set}^{\mathbf{Cell}^{\mathrm{op}}}$.

We can thus define a functor, $\mathcal{C}_K^\Delta : \mathbb{R}^d\text{-}\mathbf{Top} \to \mathbf{Set}^{\mathbf{Cell}^{\mathrm{op}}}$, mapping objects (X, f) to C_K^f, as given above, and morphisms $(X, f) \to (Y, g)$, a function-preserving map, to $C_K^f \to C_K^g$, a natural transformation [19].

3.7.2   Convergence of abstract mappers to abstract Reeb spaces

In order to compare abstract mappers, objects in the category $\mathbf{Set}^{\mathbf{Cell}^{\mathrm{op}}}$, to Reeb spaces, objects in the category $\mathbf{Set}^{\mathbf{Open}(\mathbb{R}^d)}$, we must push functors $C_K^f : \mathbf{Cell}(K)^{\mathrm{op}} \to \mathbf{Set}$ to functors $\mathbf{Open}(\mathbb{R}^d) \to \mathbf{Set}$.

Munch and Wang define the functor $\mathcal{P}_K : \mathbf{Set}^{\mathbf{Cell}^{\mathrm{op}}} \to \mathbf{Set}^{\mathbf{Open}(\mathbb{R}^d)}$ to push the abstract mapper into the category $\mathbf{Set}^{\mathbf{Open}(\mathbb{R}^d)}$. We summarize their construction and results here.

For an open set $A \subseteq \mathbb{R}^d$, define

$$K_A = \{\sigma \in K \mid \mathcal{U}_\sigma \cap A \neq \emptyset\}.$$

Remark 3.26. If I and J are open sets of $\mathbb{R}^d$ such that $I \subseteq J$, then $K_I \subseteq K_J$.

Given a functor $F : \mathbf{Cell}(K)^{\mathrm{op}} \to \mathbf{Set}$, define the functor $\mathcal{P}_K(F) : \mathbf{Open}(\mathbb{R}^d) \to \mathbf{Set}$ by

$$\mathcal{P}_K(F)(I) = \operatorname*{colim}_{\sigma \in K_I} F(\sigma),$$

for all $I \in \mathbf{Open}(\mathbb{R}^d)$.

Munch and Wang then prove the following equivalence.

Proposition 3.27 ([19]). *Let $\mathcal{F} : \mathbf{Open}(\mathbb{R}^d) \to \mathbf{Set}$ be a functor that maps an open set I to a set $\pi_0 f^{-1}\left(\bigcup_{\sigma \in K_I} \mathcal{U}_\sigma\right)$, with morphisms induced by π_0 on the inclusions. Then the functor $\mathcal{P}_K \mathcal{C}_K(X, f)$ is equivalent to $\mathcal{F}$.*

Applying this mapping, which pushes abstract mappers from the category $\mathbf{Set}^{\mathbf{Cell}^{\mathrm{op}}}$ to the category $\mathbf{Set}^{\mathbf{Open}(\mathbb{R}^d)}$, Munch and Wang give the following convergence result between the Reeb space $\mathcal{C}(X, f)$ and $\mathcal{P}_K \mathcal{C}_K(X, f)$, the abstract mapper pushed into the category $\mathbf{Set}^{\mathbf{Open}(\mathbb{R}^d)}$.

Theorem 3.28 ([19]). *Let X be a compact topological space, and let $f : X \to \mathbb{R}^d$ be a continuous multivariate function. Let $\mathcal{U}$ be a finite open cover of $\mathrm{image}(f) \subseteq \mathbb{R}^d$, and let K be the abstract simplicial complex given by the nerve of $\mathcal{U}$. Then*

$$d^{\mathcal{J}}(\mathcal{C}(X, f), \mathcal{P}_K \mathcal{C}_K(X, f)) \leq \mathrm{res}(\mathcal{U}).$$

3.7.3 Abstract mapper and hierarchical abstract mapper for point clouds

The abstract mapper stores the data of classical mapper constructions for topological spaces as a functor. This enables a formal comparison between mapper constructions and Reeb spaces. Theorem 3.28 states that as the resolution of the cover $\mathcal{U}$ goes to zero, the abstract mapper converges to the Reeb space.

When the underlying space is instead a point cloud, the classical mapper construction no longer consists simply of the nerve of a set of path-components. Instead, the classical mapper construction for point clouds is given by the nerve of the set of path-components of a Vietoris-Rips complex (Definition 3.24).

In this section, we store the data of classical mapper constructions for point clouds as a functor. This enables us to compare mapper constructions arising from a sampling or approximation of a topological space to the mapper construction, or Reeb space, of the topological space. Furthermore, we introduce the hierarchical abstract mapper, which enables the study of features captured by mapper constructions over a range of scales.

Let (X, d) be a finite metric space, such that the underlying set X is a subset of $\mathbb{R}^n$. As before, we let $f : X \to \mathbb{R}^d$ be a map into a d-dimensional parameter space and $\mathcal{U}$ be a finite open cover of the image of f. We denote the simplicial complex for the nerve of $\mathcal{U}$ by K, and for each simplex $\sigma \in K$, we denote the associated open set in $\mathbb{R}^d$ by $\mathcal{U}_\sigma = \bigcap_{\alpha \in \sigma} U_\alpha$.

Definition 3.29 (ε-expansion)**.** We define the ε-expansion of open sets on a metric space $X \subseteq \mathbb{R}^n$ as a functor

$$\Omega^\varepsilon_{\mathbb{R}^n} : \mathbf{Open}(X) \to \mathbf{Open}(\mathbb{R}^n),$$

mapping objects I of $\mathbf{Open}(X)$ to

$$I^\varepsilon := \{x \in \mathbb{R}^n : d(x, I) < \varepsilon\},$$

and mapping morphisms $I \subseteq J$ to $I^\varepsilon \subseteq J^\varepsilon$.

When the ambient space $\mathbb{R}^n$ is clear from context, we omit the subscript and simply write Ω^ε for $\Omega^\varepsilon_{\mathbb{R}^n}$.

Remark 3.30. Let I be an open set of $\mathbb{R}^d$ such that $\mathcal{U}_\sigma \cap I \neq \emptyset$. Then

$$\mathcal{U}_\sigma \subseteq I^{\mathrm{res}(\mathcal{U})}.$$

Definition 3.31. The *abstract mapper of* (X, f) *as a functor from* $\mathbf{Cell}(K)^{\mathrm{op}}$ *for the cover* $\mathcal{U}$ *and clustering parameter* δ is a functor $\dot{C}^f_{K,\delta} : \mathbf{Cell}(K)^{\mathrm{op}} \to \mathbf{Set}$ given by

$$\dot{C}^f_{K,\delta}(\sigma) := \pi_0 \Omega^\delta f^{-1}(\mathcal{U}_\sigma).$$

For $\sigma \leq \tau$, we let the map $\dot{C}^f_{K,\delta}(\tau) \to \dot{C}^f_{K,\delta}(\sigma)$ be the set map induced by the inclusion

$$\pi_0 \Omega^\delta f^{-1}(\mathcal{U}_\tau) \subseteq \pi_0 \Omega^\delta f^{-1}(\mathcal{U}_\sigma).$$

To define an interleaving distance between mapper constructions over point clouds, we will push objects in the category $\mathbf{Set}^{\mathbf{Cell}^{\mathrm{op}}}$ to the category $\mathbf{Set}^{\mathbf{Open}(\mathbb{R}^{d+1})}$.

Remark 3.32. If I and J are open sets in $\mathbb{R}^d$ such that $I \subseteq J$, then

$$\bigcup_{\sigma \in K_I} \mathcal{U}_\sigma \subseteq \bigcup_{\tau \in K_J} \mathcal{U}_\tau.$$

Definition 3.33. The *abstract mapper of* (X, f) *as a functor from* $\mathbf{Open}(\mathbb{R}^d)$ *for the cover* $\mathcal{U}$ *and clustering parameter* δ is a functor $\mathring{C}^f_{K,\delta} : \mathbf{Open}(\mathbb{R}^d) \to \mathbf{Set}$ given by

$$\mathring{C}^f_{K,\delta}(I) := \pi_0 \Omega^\delta f^{-1}\left(\bigcup_{\sigma \in K_I} \mathcal{U}_\sigma\right).$$

For $I \subseteq J$, we let the map $\mathring{C}^f_{K,\delta}(I) \to \mathring{C}^f_{K,\delta}(J)$ be the set map induced by the inclusion

$$\pi_0 \Omega^\delta f^{-1}\left(\bigcup_{\sigma \in K_I} \mathcal{U}_\sigma\right) \subseteq \pi_0 \Omega^\delta f^{-1}\left(\bigcup_{\tau \in K_J} \mathcal{U}_\tau\right)$$

implied by Lemma 3.32.

Definition 3.34. The *hierarchical abstract mapper of* (X, f) *from* $\mathbf{Open}(\mathbb{R}^d)$ is the collection $\mathring{\mathfrak{C}}^f_K = \left\{\mathring{C}^f_{K,\delta}\right\}_{\delta \geq 0}$ of functors $\mathring{C}^f_{K,\delta} : \mathbf{Open}(\mathbb{R}^d) \to \mathbf{Set}$.

Definition 3.35 (ε-interleaving)**.** Let X and Y be finite sets of the metric space $(\mathbb{R}^n, d)$. Let $f : X \to \mathbb{R}^d$ and $g : Y \to \mathbb{R}^d$ be maps into d-dimensional parameter spaces. Denote $f = (X, f)$ and $g = (Y, g)$. An ε-*interleaving* between the hierarchical abstract mappers $\mathring{\mathfrak{C}}^f_K$, $\mathring{\mathfrak{C}}^g_{K,\delta}$ is a pair of two-parameter families of maps

$$\phi_{I,\delta} : \mathring{C}^f_{K,\delta}(I) \to \mathring{C}^g_{K,\delta+\varepsilon}(I^\varepsilon), \quad \psi_{I,\delta} : \mathring{C}^g_{K,\delta}(I) \to \mathring{C}^f_{K,\delta+\varepsilon}(I^\varepsilon)$$

that are natural with respect to inclusions $I \subseteq J$ and such that the following diagrams commute for every $\delta \geq 0$:

where the maps $\mathring{C}^f_{K,\delta}(I) \to \mathring{C}^f_{K,\delta+2\varepsilon}(I^{2\varepsilon})$, $\mathring{C}^g_{K,\delta}(I) \to \mathring{C}^g_{K,\delta+2\varepsilon}(I^{2\varepsilon})$ are induced by the inclusions

$$\Omega^\delta f^{-1}\left(\bigcup_{\sigma \in K_I} \mathcal{U}_\sigma\right) \subseteq \Omega^\delta f^{-1}\left(\bigcup_{\sigma \in K_{I^{2\varepsilon}}} \mathcal{U}_\sigma\right) \subseteq \Omega^{\delta+2\varepsilon} f^{-1}\left(\bigcup_{\sigma \in K_{I^{2\varepsilon}}} \mathcal{U}_\sigma\right)$$

$$\Omega^\delta g^{-1}\left(\bigcup_{\sigma \in K_I} \mathcal{U}_\sigma\right) \subseteq \Omega^\delta g^{-1}\left(\bigcup_{\sigma \in K_{I^{2\varepsilon}}} \mathcal{U}_\sigma\right) \subseteq \Omega^{\delta+2\varepsilon} g^{-1}\left(\bigcup_{\sigma \in K_{I^{2\varepsilon}}} \mathcal{U}_\sigma\right)$$

When there exists an ε-interleaving between two hierarchical abstract mappers $\mathring{\mathfrak{C}}^f_K$ and $\mathring{\mathfrak{C}}^g_K$, we say that $\mathring{\mathfrak{C}}^f_K$, $\mathring{\mathfrak{C}}^g_K$ are ε-interleaved.

Definition 3.36 (interleaving distance)**.** The *interleaving distance* between two hierarchical abstract mappers $\mathring{\mathfrak{C}}^f_K$ and $\mathring{\mathfrak{C}}^g_K$ is given by

$$d^{\mathfrak{I}}(\mathring{\mathfrak{C}}^f_K, \mathring{\mathfrak{C}}^g_K) = \inf\left\{\varepsilon \mid \mathring{\mathfrak{C}}^f_K, \mathring{\mathfrak{C}}^g_K \text{ are } \varepsilon\text{-interleaved}\right\}.$$
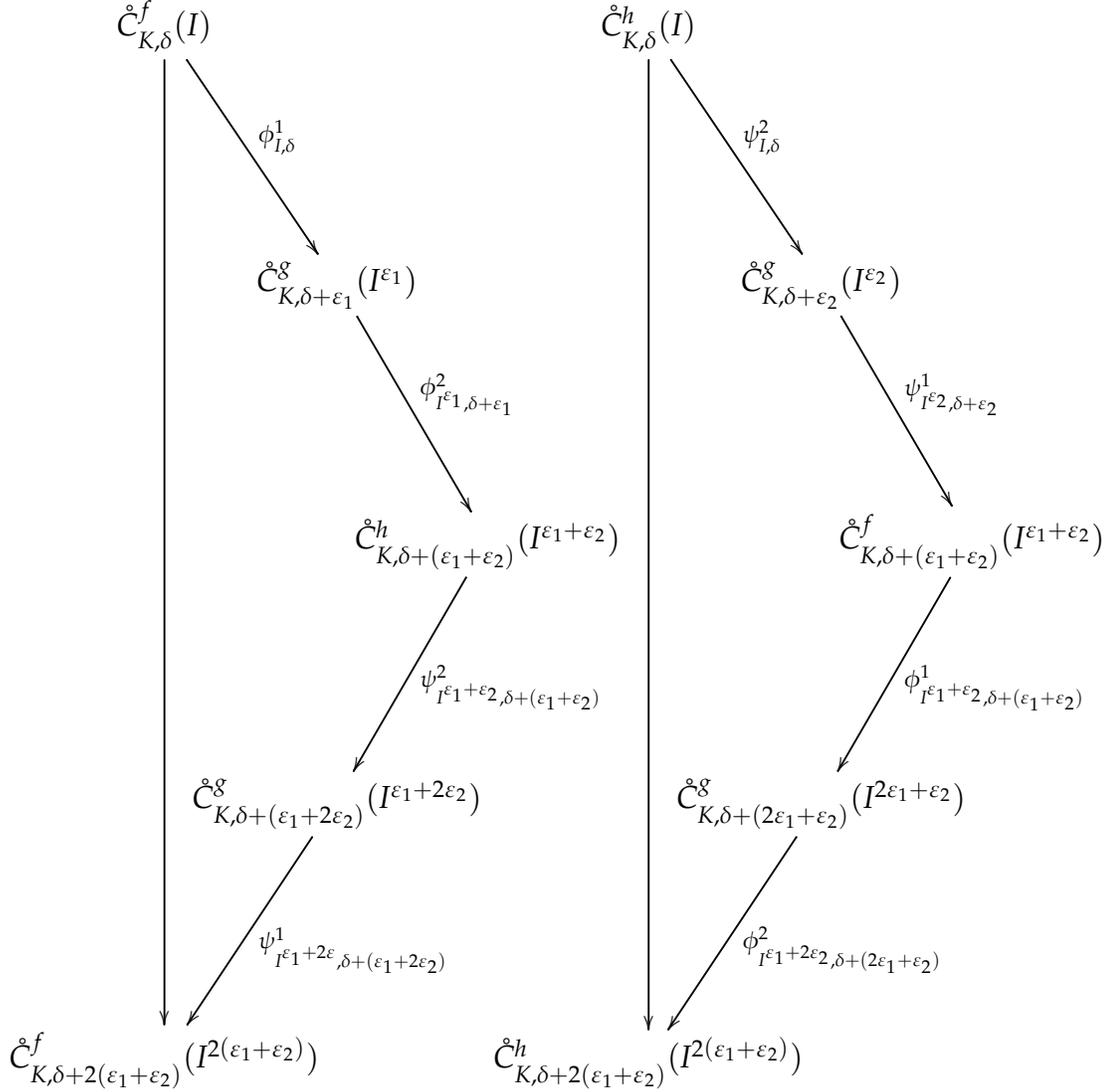
Proposition 3.37. *The interleaving distance* $d^{\mathfrak{I}}(\cdot, \cdot)$ *defines a pseudometric on the space of hierarchical abstract mappers as functors* **Open**$(\mathbb{R}^d) \to$ **Set***.*

Proof.

Let $\phi^1_{I,\delta}, \psi^1_{I,\delta}$ define an ε_1-interleaving between $\mathring{\mathfrak{C}}^f_K, \mathring{\mathfrak{C}}^g_K$, and let $\phi^2_{I,\delta}, \psi^2_{I,\delta}$ define an ε_2-interleaving between $\mathring{\mathfrak{C}}^g_K, \mathring{\mathfrak{C}}^h_K$. Then the compositions

$$\phi_{I,\delta} = \phi^2_{I^{\varepsilon_1},\delta+\varepsilon_1} \circ \phi^1_{I,\delta}, \quad \psi_{I,\delta} = \psi^1_{I^{\varepsilon_2},\delta+\varepsilon_2} \circ \psi^2_{I,\delta}$$

define an $(\varepsilon_1 + \varepsilon_2)$-interleaving between $\mathring{\mathfrak{C}}^f_K, \mathring{\mathfrak{C}}^h_K$:

$$\mathring{C}^f_{K,\delta}(I) \qquad\qquad\qquad \mathring{C}^h_{K,\delta}(I)$$

$$\phi^1_{I,\delta} \qquad\qquad\qquad \psi^2_{I,\delta}$$

$$\mathring{C}^g_{K,\delta+\varepsilon_1}(I^{\varepsilon_1}) \qquad\qquad\qquad \mathring{C}^g_{K,\delta+\varepsilon_2}(I^{\varepsilon_2})$$

$$\phi^2_{I^{\varepsilon_1},\delta+\varepsilon_1} \qquad\qquad\qquad \psi^1_{I^{\varepsilon_2},\delta+\varepsilon_2}$$

$$\mathring{C}^h_{K,\delta+(\varepsilon_1+\varepsilon_2)}(I^{\varepsilon_1+\varepsilon_2}) \qquad\qquad\qquad \mathring{C}^f_{K,\delta+(\varepsilon_1+\varepsilon_2)}(I^{\varepsilon_1+\varepsilon_2})$$

$$\psi^2_{I^{\varepsilon_1+\varepsilon_2},\delta+(\varepsilon_1+\varepsilon_2)} \qquad\qquad\qquad \phi^1_{I^{\varepsilon_1+\varepsilon_2},\delta+(\varepsilon_1+\varepsilon_2)}$$

$$\mathring{C}^g_{K,\delta+(\varepsilon_1+2\varepsilon_2)}(I^{\varepsilon_1+2\varepsilon_2}) \qquad\qquad\qquad \mathring{C}^g_{K,\delta+(2\varepsilon_1+\varepsilon_2)}(I^{2\varepsilon_1+\varepsilon_2})$$

$$\psi^1_{I^{\varepsilon_1+2\varepsilon},\delta+(\varepsilon_1+2\varepsilon_2)} \qquad\qquad\qquad \phi^2_{I^{\varepsilon_1+2\varepsilon_2},\delta+(2\varepsilon_1+\varepsilon_2)}$$

$$\mathring{C}^f_{K,\delta+2(\varepsilon_1+\varepsilon_2)}(I^{2(\varepsilon_1+\varepsilon_2)}) \qquad \mathring{C}^h_{K,\delta+2(\varepsilon_1+\varepsilon_2)}(I^{2(\varepsilon_1+\varepsilon_2)})$$

Thus, the triangle inequality holds for the interleaving distance. It follows from definition that the interleaving distance is symmetric and $d^{\mathcal{I}}(\mathring{\mathfrak{C}}^f_K, \mathring{\mathfrak{C}}^f_K) = 0$.

$\square$

Theorem 3.38. *Let (X, d) be a finite metric space. Let $f, g : X \to \mathbb{R}^d$ be filter functions into a d-dimensional parameter space $\mathbb{R}^d$. Then as $\mathrm{res}(\mathcal{U}) \to 0$, where $\mathcal{U}$ is a finite open cover for $\mathrm{image}(f) \cup \mathrm{image}(g)$, we have that*

$$d^{\mathcal{J}}(\mathring{C}^f_{K,\delta}, \mathring{C}^g_{K,\delta}) \leq \|f - g\|_\infty$$

for all $\delta \geq 0$.

Proof.

Let $\varepsilon > \|f - g\|_\infty$. Let $\mathcal{U}$ be a finite open cover of $\mathrm{image}(f) \cup \mathrm{image}(g)$ such that

$$\mathrm{res}(\mathcal{U}) + \|f - g\|_\infty \leq \varepsilon.$$

We wish to show that

$$f^{-1}\left(\bigcup_{\sigma \in K_I} \mathcal{U}_\sigma\right) \subseteq g^{-1}\left(\bigcup_{\sigma \in K_{I^\varepsilon}} \mathcal{U}_\sigma\right), \tag{3.3}$$

for all open $I \subseteq \mathbb{R}^d$.

Consider $x \in f^{-1}\left(\bigcup_{\sigma \in K_I} \mathcal{U}_\sigma\right)$. Then there exists σ such that $\mathcal{U}_\sigma \cap I \neq \emptyset$ and $f(x) \in \mathcal{U}_\sigma$. By Remark 3.30, we have that $f(x) \in I^{\mathrm{res}(\mathcal{U})}$. Since

$$|f(x) - g(x)| \leq \|f - g\|_\infty < \mathrm{res}(\mathcal{U}) + \|f - g\|_\infty \leq \varepsilon,$$

we have that $g(x) \in I^{\mathrm{res}(\mathcal{U})}$. Since $\mathcal{U}$ is a cover of $\mathrm{image}(g)$, then $x \in g^{-1}\left(\bigcup_{\sigma \in K_I} \mathcal{U}_\sigma\right)$. So the inclusion (3.3) holds.

Similarly, we have that

$$g^{-1}\left(\bigcup_{\sigma \in K_I} \mathcal{U}_\sigma\right) \subseteq f^{-1}\left(\bigcup_{\sigma \in K_{I^\varepsilon}} \mathcal{U}_\sigma\right). \tag{3.4}$$

The inclusions 3.3 and 3.4 imply that for all $\delta \geq 0$, we have the following

diagram of inclusions

$$\Omega^\delta f^{-1}\left(\bigcup_{\sigma\in K_I}\mathcal{U}_\sigma\right)$$

$$\Omega^\delta g^{-1}\left(\bigcup_{\sigma\in K_{I^\varepsilon}}\mathcal{U}_\sigma\right)\longrightarrow\Omega^{\delta+\varepsilon}g^{-1}\left(\bigcup_{\sigma\in K_{I^\varepsilon}}\mathcal{U}_\sigma\right)$$

$$\Omega^\delta f^{-1}\left(\bigcup_{\sigma\in K_{I^{2\varepsilon}}}\mathcal{U}_\sigma\right)\longrightarrow\Omega^{\delta+\varepsilon}f^{-1}\left(\bigcup_{\sigma\in K_{I^{2\varepsilon}}}\mathcal{U}_\sigma\right)$$

$$\Omega^{\delta+\varepsilon}f^{-1}\left(\bigcup_{\sigma\in K_{I^{2\varepsilon}}}\mathcal{U}_\sigma\right)\longrightarrow\Omega^{\delta+2\varepsilon}f^{-1}\left(\bigcup_{\sigma\in K_{I^{2\varepsilon}}}\mathcal{U}_\sigma\right)$$

and

$$\Omega^\delta g^{-1}\left(\bigcup_{\sigma\in K_I}\mathcal{U}_\sigma\right)$$

$$\Omega^\delta f^{-1}\left(\bigcup_{\sigma\in K_{I^\varepsilon}}\mathcal{U}_\sigma\right)\longrightarrow\Omega^{\delta+\varepsilon}f^{-1}\left(\bigcup_{\sigma\in K_{I^\varepsilon}}\mathcal{U}_\sigma\right)$$

$$\Omega^\delta g^{-1}\left(\bigcup_{\sigma\in K_{I^{2\varepsilon}}}\mathcal{U}_\sigma\right)\longrightarrow\Omega^{\delta+\varepsilon}g^{-1}\left(\bigcup_{\sigma\in K_{I^{2\varepsilon}}}\mathcal{U}_\sigma\right)$$

$$\Omega^{\delta+\varepsilon}g^{-1}\left(\bigcup_{\sigma\in K_{I^{2\varepsilon}}}\mathcal{U}_\sigma\right)\longrightarrow\Omega^{\delta+2\varepsilon}g^{-1}\left(\bigcup_{\sigma\in K_{I^{2\varepsilon}}}\mathcal{U}_\sigma\right)$$

Applying π_0, we obtain an ε-interleaving given by the outermost arrows of the commutative diagrams:

$$\phi_{I,\delta}:\mathring{C}^f_{K,\delta}(I)\to\mathring{C}^g_{K,\delta+\varepsilon}(I^\varepsilon),\quad\psi_{I,\delta}:\mathring{C}^g_{K,\delta}(I)\to\mathring{C}^f_{K,\delta+\varepsilon}(I^\varepsilon).$$

Thus,

$$d^{\mathcal{I}}(\mathring{\mathfrak{C}}^f_K,\mathring{\mathfrak{C}}^g_K)\leq\|f-g\|_\infty$$

for all $\delta\geq 0$.

$\square$

Remark 3.39. The uppermost triangle in each commutative diagram shows that whenever $\|f-g\|_\infty<\varepsilon$, the abstract mappers $\mathring{C}^f_{K,\delta}$ and $\mathring{C}^g_{K,\delta}$ for each fixed $\delta\geq 0$ are ε-interleaved with respect to inclusion. This is consistent with the notion of ε-interleaving as defined by Chazal et al. [88].

Definition 3.40 (Hausdorff distance). Let (Z,d) be a compact metric space. The *Hausdorff distance* between any two compact subsets $X, Y \subseteq Z$ is

$$d^{\mathcal{H}}(X,Y) := \max\left(\max_{x \in X} \min_{y \in Y} d(x,y), \max_{y \in Y} \min_{x \in X} d(x,y)\right).$$

Theorem 3.41. *Let X and Y be compact subsets of a metric space (Z,d). Let $f : X \cup Y \to \mathbb{R}^d$ be a Lipschitz continuous filter function into $\mathbb{R}^d$, with Lipschitz constant $L \in \mathbb{R}$. Denote the restrictions of f to X and Y by $f_X : X \to \mathbb{R}^d$ and $f_Y : Y \to \mathbb{R}^d$, respectively. Then as $\mathrm{res}(\mathcal{U}) \to 0$, where $\mathcal{U}$ is an open cover for $\mathrm{image}(f) \cup \mathrm{image}(g)$, we have that*

$$d^{\mathcal{I}}(\mathring{\mathfrak{C}}_K^{f_X}, \mathring{\mathfrak{C}}_K^{f_Y}) \leq \max\left(d^{\mathcal{H}}(X,Y),\ L \cdot d^{\mathcal{H}}(X,Y)\right),$$

where $d^{\mathcal{H}}(X,Y)$ is the Hausdorff distance between X and Y.

Proof.

Let $\varepsilon > \max\left(d^{\mathcal{H}}(X,Y),\ L \cdot d^{\mathcal{H}}(X,Y)\right)$. Let $\mathcal{U}$ be a finite open cover of $\mathrm{image}(f)$ such that

$$\mathrm{res}(\mathcal{U}) + L \cdot d^{\mathcal{H}}(X,H) \leq \varepsilon.$$

We wish to show that

$$f_X^{-1}\left(\bigcup_{\sigma \in K_I} \mathcal{U}_\sigma\right) \subseteq \Omega^\varepsilon f_Y^{-1}\left(\bigcup_{\sigma \in K_{I^\varepsilon}} \mathcal{U}_\sigma\right). \tag{3.5}$$

Consider $x \in f_X^{-1}\left(\bigcup_{\sigma \in K_I} \mathcal{U}_\sigma\right)$. Then there exists a simplex σ such that $\mathcal{U}_\sigma \cap I \neq \emptyset$ and $f_X(x) \in \mathcal{U}_\sigma \subseteq I^{\mathrm{res}(\mathcal{U})}$. By definition of Hausdorff distance, there exists $y \in Y$ such that $|x - y| < d^{\mathcal{H}}(X,Y) < \varepsilon$. Furthermore, by Lipschitz continuity, we have that

$$|f(x) - f(y)| \leq L \cdot |x - y| \leq L \cdot d^{\mathcal{H}}(X,Y).$$

So $f(y) \in I^{\mathrm{res}(\mathcal{U}) + L \cdot d^{\mathcal{H}}(X,Y)} \subseteq I^\varepsilon$ and $y \in f_Y^{-1}\left(\bigcup_{\sigma \in K_{I^\varepsilon}} \mathcal{U}_\sigma\right)$. Hence,

$$x \in \Omega^\varepsilon f_Y^{-1}\left(\bigcup_{\sigma \in K_{I^\varepsilon}} \mathcal{U}_\sigma\right).$$

So the inclusion (3.5) holds. It follows that

$$\Omega^\delta f_X^{-1}\left(\bigcup_{\sigma\in K_I}\mathcal{U}_\sigma\right)\subseteq\Omega^{\delta+\varepsilon}f_Y^{-1}\left(\bigcup_{\sigma\in K_{I^\varepsilon}}\mathcal{U}_\sigma\right). \tag{3.6}$$

Similarly,

$$\Omega^\delta f_Y^{-1}\left(\bigcup_{\sigma\in K_I}\mathcal{U}_\sigma\right)\subseteq\Omega^{\delta+\varepsilon}f_X^{-1}\left(\bigcup_{\sigma\in K_{I^\varepsilon}}\mathcal{U}_\sigma\right). \tag{3.7}$$

The inclusions (3.6) and (3.7) imply that for all $\delta\geq 0$, we have

$$\pi_0\Omega^\delta f_X^{-1}\left(\bigcup_{\sigma\in K_I}\mathcal{U}_\sigma\right)\subseteq\pi_0\Omega^{\delta+\varepsilon}f_Y^{-1}\left(\bigcup_{\sigma\in K_{I^\varepsilon}}\mathcal{U}_\sigma\right),$$

$$\pi_0\Omega^\delta f_Y^{-1}\left(\bigcup_{\sigma\in K_I}\mathcal{U}_\sigma\right)\subseteq\pi_0\Omega^{\delta+\varepsilon}f_X^{-1}\left(\bigcup_{\sigma\in K_{I^\varepsilon}}\mathcal{U}_\sigma\right),$$

which induce an ε-interleaving

$$\phi_{I,\delta}:\mathring{C}^{f_X}_{K,\delta}(I)\to\mathring{C}^{f_Y}_{K,\delta+\varepsilon}(I^\varepsilon),\quad\psi_{I,\delta}:\mathring{C}^{f_Y}_{K,\delta}(I)\to\mathring{C}^{f_X}_{K,\delta+\varepsilon}(I^\varepsilon).$$

So

$$d^{\mathfrak{I}}(\mathring{\mathfrak{C}}^{f_X}_K,\mathring{\mathfrak{C}}^{f_Y}_K)\leq\max\left(d^{\mathcal{H}}(X,Y),\,L\cdot d^{\mathcal{H}}(X,Y)\right).$$

$$\square$$

3.7.4   Correspondence between single-linkage hierarchical clustering and hierarchical mapper constructions

In this section, we show that partitions of a finite metric space (X,d) obtained as an output of single-linkage clustering can be obtained as mapper constructions over X.

Preliminaries

For a finite set X, let $\mathcal{P}(X)$ denote set of all partitions of X.

Definition 3.42 (persistent set [86]). A *persistent set* is a pair (X,θ_X), given by a finite set X and a function $\theta_X:[0,\infty)\to\mathcal{P}(X)$ from the positive real numbers to the set of all partitions of X, such that the following properties hold:

1. if $r \leq s$, then $\theta_X(r)$ refines $\theta_X(s)$;

2. for any r, there exists $\varepsilon > 0$ such that $\theta_X(r') = \theta_X(r)$ for all $r' \in [r, r + \varepsilon]$.

Definition 3.43 (dendrogram)**.** Let (X, θ_X) be a persistent pair. If there exists $t > 0$ such that $\theta_X(t)$ consists of the single block partition for all $r \geq t$, then (X, θ_X) is a *dendrogram*.

Example 3.44 (single-linkage hierarchical clustering [21, 86])**.** For each finite metric space (X, d_X), we can associate a persistent pair (X, θ_X), where for each $r \geq 0$, blocks of the partition $\theta_X(r)$ consist of the equivalence classes of $\sim_r$. The pair (X, θ_X) is the *dendrogram* for single-linkage hierarchical clustering over (X, d_X).

Let (X, d) be a finite metric space, $X \subseteq \mathbb{R}^n$, and let $\chi_X : X \to \mathbb{R}$ be the constant function $\chi_X(x) = 0$, with finite open cover $\mathcal{U}$ of image(χ_X).

For every open set $I \subseteq \mathbb{R}$ containing zero,

$$\Omega^\delta \chi_X^{-1}\left(\bigcup_{\sigma \in K_I} \mathcal{U}_\sigma\right) = \Omega^\delta X$$

consists of path-components such that $x, x' \in X$ belong to the same path-component if and only if there exists a sequence of points $x_0, \cdots, x_k \in X$ such that $x_0 = x$, $x_k = x'$, and $d(x_i, x_{i+1}) \leq \delta$. That is, path-components of $\Omega^\delta X$ are exactly the $\sim_\delta$ equivalence classes of X.

Thus, there is a correspondence between the output of single-linkage clustering at the clustering resolution δ, denoted $\theta_X(\delta)$, and the abstract mapper $\mathring{C}^{\chi_X}_{K,\delta}$, which we denote $\mathring{C}^X_{K,\delta}$ for simplicity.

Furthermore, if we define $\Xi^X_K : [0, \infty) \to \mathbf{Set}^{\mathbf{Open}(\mathbb{R}^d)}$ by $\Xi^X_K(r) = C^X_{K,\delta}$, we have a correspondence between single-linkage hierarchical clustering and hierarchical abstract mappers.

Interleaving distance and Hausdorff distance

Proposition 3.45. *Let X and Y be finite sets of $\mathbb{R}^n$. Let $\chi_X : X \to \mathbb{R}$ and $\chi_Y : Y \to \mathbb{R}$ be the constant zero functions on X and Y. Denote their hierarchical mapper constructions by $\mathring{\mathfrak{C}}^X_K$ and $\mathring{\mathfrak{C}}^Y_{K,\delta}$, respectively. Then*

$$d^{\mathfrak{I}}(\mathring{\mathfrak{C}}^X_K, \mathring{\mathfrak{C}}^Y_K) \leq d^{\mathcal{H}}(X, Y),$$

for all δ ≥ 0.

Proof.

Constant functions are 0-Lipschitz. So by Proposition 3.41, the inequality holds.

$\square$

Interleaving distance and Gromov-Hausdorff distance

Definition 3.46. For a finite metric space (X, d), define $u : X \times X \to \mathbb{R}$ by

$$u(x, x') = \min \{r \geq 0 \mid x \sim_r x'\}.$$

Remark 3.47. $u : X \times X \to \mathbb{R}$ is in fact an *ultrametric*:

1. $u(x, x') \geq 0$ for all $x, x' \in X$;

2. $u(x, x') = 0$ iff $x = x'$;

3. $u(x, x') = u(x', x)$;

4. $u(x, z) \leq max(u(x, y), u(y, z))$.

Carlsson and Mémoli [21] give a bijection between dendrograms and ultrametric spaces.

Repeatedly applying Condition 4, known as the *strong inequality* or *ultrametric inequality*, we obtain

$$u_D(x_1, x_k) \leq \max(u_D(x_1, x_2), u_D(x_2, x_3), \cdots, u_D(x_{k-1}, x_k)). \qquad (3.8)$$

Definition 3.48 (distortion [21])**.** Let X, Y be finite sets. Define the *distortion* of $f : X \to Y$ by

$$\mathrm{dis}(f) := \max_{x, x' \in X} \left| u_X(x, x') - u_Y(f(x), f(x')) \right|.$$

Definition 3.49 (joint distortion [21])**.** Let X, Y be finite sets. Define the *joint distortion* of $f : X \to Y$ and $g : Y \to X$ by

$$\mathrm{dis}(f, g) := \max_{x \in X, y \in Y} \left| u_X(x, g(y)) - u_Y(f(x), y) \right|.$$

Definition 3.50 (Gromov-Hausdorff distance [21]). Let X and Y be finite sets. Define the *Gromov-Hausdorff* distance between the dendrograms (X, u_X) and (Y, u_Y) for single-linkage hierarchical clustering over X and Y, respectively, as

$$d^{\mathcal{GH}}(X, Y) := \frac{1}{2} \min_{f, g} \max \left(\mathrm{dis}(f), \mathrm{dis}(g), \mathrm{dis}(f, g) \right),$$

where $f : X \to Y$ and $g : Y \to X$ are functions between X and Y.

Proposition 3.51. *Let X and Y be finite sets in $(\mathbb{R}^n, d)$. Then*

$$d^{\mathcal{J}}(\mathring{\mathfrak{C}}_K^X, \mathring{\mathsf{C}}_{\mathfrak{R}}^Y) \leq 2 \cdot d^{\mathcal{GH}}((X, u_X), (Y, u_Y)),$$

where $d^{\mathcal{GH}}((X, u_X), (Y, u_Y))$ is the Gromov-Hausdorff distance between the ultra-metric spaces (X, u_X) and (U, u_Y).

Proof.

Suppose $\varepsilon \geq 2 \cdot d^{\mathcal{GH}}((X, u_X), (Y, u_Y))$. Then there exist maps $f : X \to Y$ and $g : Y \to X$ such that

$$\max(\mathrm{dis}(f), \mathrm{dis}(g), \mathrm{dis}(f, g)) \leq \varepsilon. \tag{3.9}$$

Since the joint distortion of f and g, $\mathrm{dis}(\phi, \psi)$, is bounded above by ε, then $\left| u_X(x, g(y)) - u_Y(f(x), y) \right| \leq \varepsilon$ for all $x \in X$ and $y \in Y$. In particular, for $y = \phi(x)$, we have that

$$u_X(x, g(f(x))) = \left| u_X(x, g(f(x))) - u_Y(f(x), f(x)) \right| \leq \varepsilon. \tag{3.10}$$

Similarly,

$$u_Y(f(g(y)), y) \leq \varepsilon. \tag{3.11}$$

It is sufficient to show that for all $\delta \geq 0$, there exists a pair of maps

$$\phi_\delta : \pi_0 \Omega^\delta X \to \pi_0 \Omega^{\delta + \varepsilon} Y, \quad \psi_\delta : \pi_0 \Omega^\delta Y \to \pi_0 \Omega^{\delta + \varepsilon} X$$

such that the following diagrams commute:

We define a map $f^\delta : \Omega^\delta X \to \Omega^{\delta+\varepsilon} Y$ as follows: for each path-component X_i of $\Omega^\delta X$,

- if X_i contains at least one point of the image $g : Y \to X$, then we choose an arbitrary $x_i \in X_i \cap \mathrm{image}(g)$, and we map every point of the path-component X_i to $f(x_i)$;

- if X_i intersects trivially with the image of g, then we pick an arbitrary representative $x_i \in X_i \cap X$, and we map every point of the path-component X_i to $f(x_i)$.

We define an analogous map $g^\delta : \Omega^\delta Y \to \Omega^{\delta+\varepsilon}$.

Since the image of each path-component of $\Omega^\delta X$ and $\Omega^\delta Y$ under f^δ and g^δ, respectively, is a single point, then f^δ and g^δ induce well-defined maps $\phi_\delta : \pi_0 \Omega^\delta X \to \pi_0 \Omega^{\delta+\varepsilon} Y$ and $\psi_\delta : \pi_0 \Omega^\delta Y \to \pi_0 \Omega^{\delta+\varepsilon} X$.

We now show that the diagrams commute.

If for all path-components X_i of $\Omega^\delta X$, we have that $u_Y(f(x), f(x')) \le \delta + \varepsilon$ for all $x, x' \in X_i$, then $f : X \to Y$ induces a well-defined map $f^\delta : \Omega^\delta X \to \Omega^{\delta+\varepsilon} Y$. And if, additionally, each path-component of $\Omega^{\delta+\varepsilon} Y$ contains at most one point of the image of $f : X \to Y$, then commutativity of the diagram on the left follows from inequality 3.10, for any choice of $g^{\delta+\varepsilon} : \Omega^{\delta+\varepsilon} Y \to \Omega^{\delta+2\varepsilon}$, as described above.

So suppose a path-component X_i of $\Omega^\delta X$ contains $x, x' \in X$, thus possibly requiring a (non-trivial) choice for the image of X_i under f^δ. Suppose also that there exists $y = f(x'') \in \mathrm{image}(f)$ such that y and $f(x')$ belong to the same component of $\Omega^{\delta+\varepsilon}$, thus possibly requiring a (non-trivial) choice for the image of $f(x')$ under $g^{\delta+\varepsilon}$.

Note that we do not require that $x \ne x'$ or that $f(x') \ne f(x'')$. Thus, in order to prove that the diagram on the left commutes, it is sufficient to show that

$u_X(x, g(f(x''))) \leq \delta + 2\varepsilon$:

$u_X(x, g(f(x'')))$

$\leq \max\left(u_X(x, g(f(x))),\ u_X(g(f(x)), g(f(x'')))\right)$, by the ultrametric inequality

$\leq \max(\varepsilon, u_X(g(f(x)), g(f(x''))))$, by inequality (3.10)

$\leq \max(\varepsilon, u_Y(f(x), f(x'')) + \varepsilon)$, since $\text{dis}(g) \leq \varepsilon$ by (3.9)

$\leq \max(\varepsilon, \max(u_Y(f(x), f(x')), u_Y(f(x'), f(x''))) + \varepsilon)$

$\leq \max(\varepsilon, \max(\delta + \varepsilon, u_Y(f(x'), f(x''))) + \varepsilon)$, since $\text{dis}(f) \leq \varepsilon$, $u_X(x, x') \leq \delta$

$\leq \max(\varepsilon, \max(\delta + \varepsilon, \delta + \varepsilon) + \varepsilon)$, since $f(x') \sim_{\delta+\varepsilon} f(x'')$

$\leq \max(\varepsilon, \delta + 2\varepsilon)$

$\leq \delta + 2\varepsilon$

Similarly, we can show that the diagram on the right also commutes. Thus, $\phi_\delta : \pi_0 \Omega^\delta X \to \pi_0 \Omega^{\delta+\varepsilon} Y$ and $\psi_\delta : \pi_0 \Omega^\delta Y \to \pi_0 \Omega^{\delta+\varepsilon} X$ define an ε-interleaving between $\mathring{\mathfrak{C}}^X_K$ and $\mathring{\mathfrak{C}}^Y_K$.

$\square$

3.8 Application to the analysis of geospatial sensor data

Hierarchical mapper constructions allow for the comparison of mapper constructions over a range of clustering resolutions, much like dendrograms (produced by hierarchical clustering) allow for the comparison of clusters over a range of distances.

In cases where mapper constructions contain non-trivial topological features like loops, persistent homology can be applied to hierarchical mapper constructions in order obtain the distribution of scales at which features (versus noise) exist.

In this section, we summarize the algorithm for computing mapper constructions, as given by Carlsson [76], and extend the process to provide an algorithm for computing hierarchical mapper constructions. We then apply the algorithm to a time series of daily sea surface temperatures, in order to provide a statistic, the persistence of 1-dimensional classes, to compare interannual climate variability. The persistence of the 1-dimensional classes in the hierarchical mapper construction form a clear bimodal distribution, with

a gap between the two modes (Figure 3.5, bottom graph), differentiating the El Niño/La Niña phases from the rest of the El Niño Southern Oscillation.

3.8.1 Algorithm for mapper constructions

The mapper construction described in Definition 3.24 can be obtained via the following algorithm [76]:

1. Define a filter function $f : X \rightarrow Y$ from the point cloud X.

2. Select a finite covering $\mathcal{U} = \{U_\alpha\}_{\alpha \in A}$ of Y.

3. Select a clustering threshold δ, and perform single-linkage clustering on the subsets $X_\alpha = f^{-1}(U_\alpha)$. This results in a covering of X by pairs (α, c), where $\alpha \in A$ and c is a cluster in X_α.

4. Construct a simplicial complex $\check{C}(X, f, \mathcal{U}, \delta)$ whose vertex set consists of all such pairs (α, c), and a family $\{(\alpha_0, c_0), \cdots, (\alpha_k, c_k)\}$ spans a k-simplex if and only if the corresponding clusters have a point in common.

3.8.2 Algorithm for hierarchical mapper constructions

The hierarchical mapper construction begins much like the mapper construction, but instead of selecting a single clustering threshold, we consider a finite sequence of clustering thresholds $\delta_1 < \cdots < \delta_j < \cdots < \delta_m$.

1. Define a filter function $f : X \rightarrow Y$ from the point cloud X.

2. Select a finite covering $\mathcal{U} = \{U_\alpha\}_{\alpha \in A}$ of Y.

3. For each clustering threshold δ_j, perform single-linkage clustering on the subsets $X_\alpha \times \{\delta_j\} = f^{-1}(U_\alpha) \times \{\delta_j\}$, resulting in triples (α, c, δ_j) that cover $X \times \{\delta_j\}$, with $\alpha \in A$ and c a cluster in $X_\alpha \times \{\delta_j\}$.

4. Construct a simplicial complex whose vertex set consists of all such triples (α, c, δ_j) and whose k-simplices are of one of the following forms:

   - $\{(\alpha_0, c_0, \delta_j), \cdots, (\alpha_k, c_k, \delta_j)\}$ spans a k-simplex if and only if the corresponding clusters have a point in common;
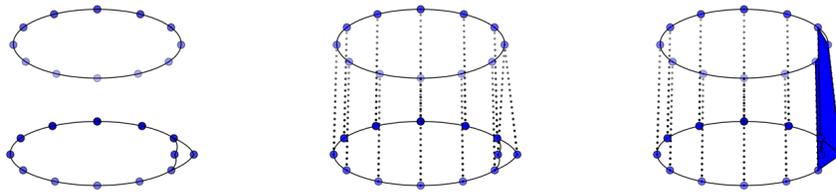
Figure 3.2: (left) Simplices within clustering thresholds. (center) 1-simplices between clustering thresholds. (right) 2-simplices between clustering thresholds.

- $\{(\alpha, c_{j,0}, \delta_j), \cdots, (\alpha, c_{j,k}, \delta_j), (\alpha, c_{j+1,0}, \delta_{j+1}), \cdots, (\alpha, c_{j+1,k'}, \delta_{j+1})\}$ spans a $(k + k' + 1)$-simplex if and only if for each $(\alpha, c_{j+1,\ell'})$, the corresponding cluster (at threshold δ_{j+1}) contains all clusters (at threshold δ_j) corresponding to $(\alpha, c_{j,\ell}), 0 \leq \ell \leq k$.
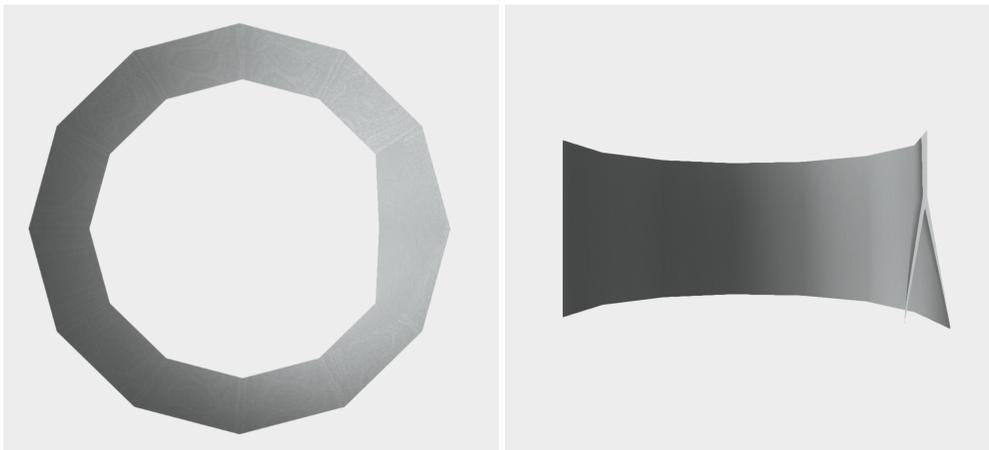

Figure 3.3: Top (left) and cross-sectional (right) views of the hierarchical mapper construction depicted in Figure 3.2.

3.8.3 Study of sea surface temperatures using persistent homology

In this example we examine sea surface temperatures as an example of a high-dimensional oscillator. The annual cycle is the dominant signal that is captured at the interannual timescale, but there is additional interannual variability introduced by oscillations such as the El Niño Southern Oscillation (ENSO). We study the magnitude of these oscillations using persistent homology.

Data

We use Advanced Very High Resolution Radiometer (AVHRR-only) data from the NOAA $1/4°$ Daily Optimum Interpolation Sea Surface Temperature analysis [89], constructed by combining observations from different platforms (e.g., satellites, ships, buoys) on a regular 720×1440 global grid, from September 1, 1981 to January 15, 2016.

Construction

We construct a hierarchical mapper complex from global sea surface temperatures. We let X be the set of data points given by $(t, \mathbf{x}_t)$, where $\mathbf{x}_t$ is a grid of global sea surface temperatures for time t.

We let the filter function $f : X \to S^1$ be given by projecting the time series onto a circle (representing the annual cycle). The filter function f maps each sample to its day of year. Let $\mathcal{U} = \{U_i\}_{i=1}^{12}$ be an open cover of S^1, such that $U_i = (r_i - \delta, s_i + \delta)$, where r_i is the first day of the ith month, s_i is the last day of the ith month, and δ is fourteen days (two weeks). (It is understood that $r_i - \delta$ and $s_i + \delta$ are computed modulo a year. e.g., December 31 plus one day is January 1.)

We then construct a hierarchical mapper complex for (X, f) using the algorithm described in Section 3.8.2. We measure the distance between two daily sea surface temperature measurements x and x' in X by the L^2-norm. We display the output in Figure 3.4.

Remark 3.52. There exist resolution parameters $\delta_1, \delta_m > 0$ such that the mapper construction given by $\bar{f}_X^*(\mathcal{U} \times \{\delta_m\})$ is a circle, and the mapper construction given by $\bar{f}_X^*(\mathcal{U} \times \{\delta_1\})$ is a spiral.

For a fixed δ such that $\delta_1 < \delta < \delta_m$, the mapper construction given by $\bar{f}_X^*(\mathcal{U} \times \{\delta\})$ may contain loops (nontrivial 1-dimensional homology classes). Each loop is a depiction of interannual variability detected at the resolution parameter δ.

Measuring persistence

We use Dionysus [90] to compute 1-dimensional persistence classes of the hierarchical mapper complex. Constructing the histogram of the persistence

of 1-dimensional classes, we find that there is a gap between 1-dimensional classes corresponding to the most salient sea surface temperature anomalies and 1-dimensional classes corresponding to the remaining sea surface temperature observations. (See Figure 3.5.)

We find that the most salient signals occur during June-July-August (JJA) and December-January-February (DJF).

3.9   Discussion and future directions

The mapper construction has found success in many applications and has stood out as a tool for extracting insights in complex, noisy, high-dimensional data. In this chapter, we have provided a practical and theoretical framework to analyze mapper constructions over a range of scales. The hierarchical mapper construction enables a systematic study of features and noise, as well as the application of statistics. Future work may identify and investigate the statistical tools that support the rigorous interpretation and analysis of mapper constructions and hierarchical mapper constructions.
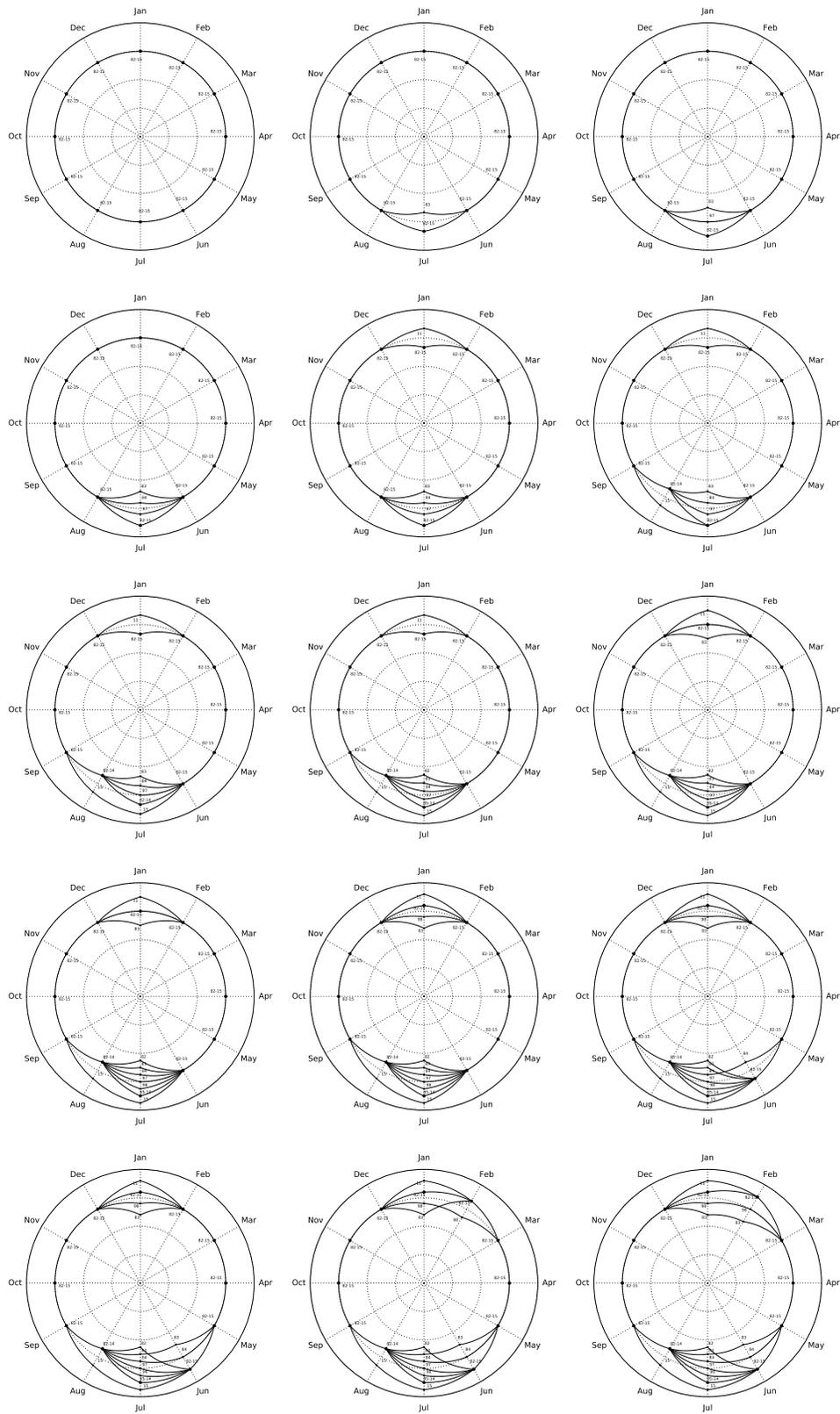
Figure 3.4: Mapper constructions for global sea surface temperatures. Loops correspond to non-trivial 1-dimensional persistence classes.
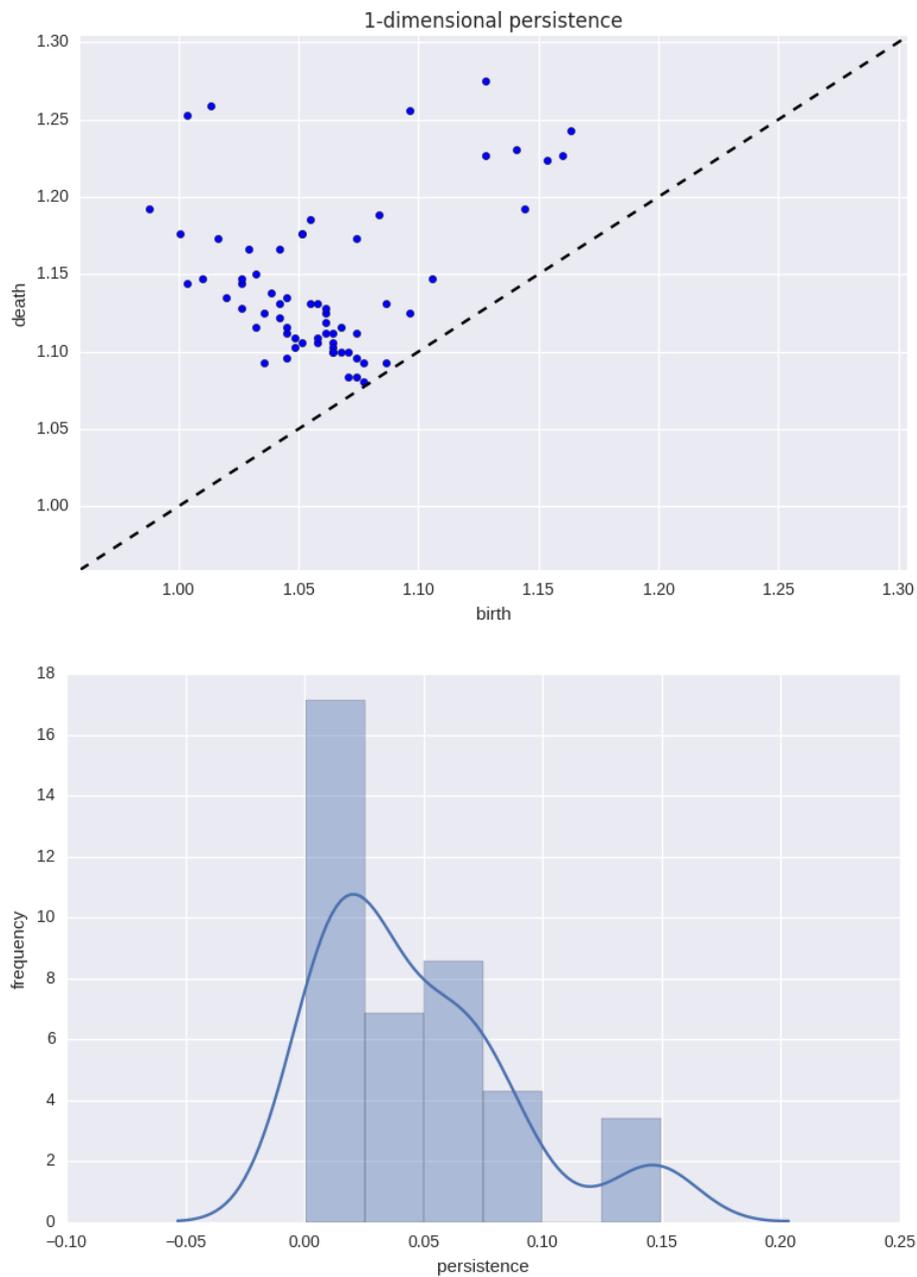
Figure 3.5: (top) Persistence diagram of 1-dimensional persistence classes, depicting the birth and death of each class. (bottom) Histogram of 1-dimensional persistence classes of hierarchical mapper construction of global sea surface temperatures.

NOMENCLATURE

B_n — the braid group on n strands — 11

$B(\alpha)(t)$ — the Burau matrix for the braid α with parameter t — 15

σ_i — the braid consisting of a single crossing given by passing the ith strand behind the $(i+1)$st strand — 10

σ_{i,η_1,η_2} — the braid that moves the group of η_1 consecutive strands starting at strand i behind the group of η_2 consecutive strands starting at strand $i + \eta_1$ — 19

$\mathbf{Cell}(K)$ — the category whose objects consist of simplices of K and whose morphisms $\sigma \to \tau$ are given by the face relation $\sigma \leq \tau$ — 59

$\mathbf{Cell}(K)^{\mathrm{op}}$ — the category whose objects consist of the simplices of K and whose morphisms $\tau \to \sigma$ are given by the face relation $\sigma \leq \tau$ — 59

$\mathbb{R}\text{-}\mathbf{Top}$ — the category whose objects consist of pairs (X, f), where X is a topological space and $f : X \to \mathbb{R}$ is a continuous map, and whose morphisms $\phi : (X, f) \to (Y, g)$ are continuous function-preserving maps $\phi : X \to Y$ — 49

$\mathbb{R}\text{-}\mathbf{Top}^{\mathrm{c}}$ — full subcategory of constructible $(X, f) \in \mathbb{R}\text{-}\mathbf{Top}$ — 50

$\mathbb{R}^d\text{-}\mathbf{Top}$ the category whose objects consist of pairs   53
(X, f), where X is a topological space and
$f : X \to \mathbb{R}^d$ is a continuous map, and whose
morphisms $\phi : (X, f) \to (Y, g)$ are continu-
ous function-preserving maps $\phi : X \to Y$

$d^{\mathcal{GH}}(X, Y)$ Gromov-Hausdorff distance between X and   70
Y

$d^{\mathcal{H}}(X, Y)$ Hausdorff distance between X and Y 67

I^ε ε–expansion of (an open) set I 61

K simplicial complex for the nerve of $\mathcal{U}$ 59

π_0 set path–components of a space   51

$\mathcal{P}(X)$ set of all partitions of a finite set X 68

$\mathrm{res}(\mathcal{U})$ resolution of a cover $\mathcal{U} = \{U_\alpha\}_{\alpha \in A}$, where   58
$\mathrm{res}(\mathcal{U}) := \sup\{\mathrm{diam}(U_\alpha) \mid U_\alpha \in \mathcal{U}\}$

$\sim_\delta$ $x \sim_\delta x'$ if and only if there is a sequence   56
$x_0, \cdots, x_k \in X$ so that $x_0 = x$, $x_k = x'$, and
$d_X(x_i, x_{i+1}) \leq \delta$

σ a simplex   44

τ a simplex   44

$\mathcal{U}$ finite open cover   56

$\mathcal{U}_\sigma$ the open set $\mathcal{U}_\sigma = \bigcap_{\alpha \in \sigma} U_\alpha$ associated to σ 59

BIBLIOGRAPHY

BIBLIOGRAPHY

1. Stillwell, J. *Mathematics and Its History* 3rd ed. ISBN: 978-1-4419-6052-8. doi:`10.1007/978-1-4419-6053-5` (Springer-Verlag New York, 2010).

2. Borel, A. in Leray, J. *Selected Papers - Oeuvres Scientifiques. Topology and Fixed Point Theorems* with an intro. by Borel, A. 3 vols. (Springer-Verlag Berlin Heidelberg, 1998). ISBN: 978-3-642-41847-1.

3. Meyer, M., Desbrun, M., Schröder, P. & Barr, A. in *Visualization and Mathematics III* (eds Hege, H.-C. & Polthier, K.) 35–57 (Springer Berlin Heidelberg, 2003). ISBN: 978-3-642-05682-6. doi:`10.1007/978-3-662-05105-4_2`.

4. *Discrete Differential Geometry* 1st ed. (eds Bobenko, A. I., Schröder, P., Sullivan, J. M. & Ziegler, G. M.) ISBN: 978-3-7643-8620-7. doi:`10.1007/978-3-7643-8621-4` (Birkhäuser Basel, 2008).

5. Singh, G., Mémoli, F. & Carlsson, G. Topological Methods for the Analysis of High Dimensional Data Sets and 3D Object Recognition. *Eurographics Symposium on Point-Based Graphics* (2007).

6. Nicolau, M., Levine, A. & Carlsson, G. Topology based data analysis identifies a subgroup of breast cancers with a unique mutational profile and excellent survival. *Proceedings of the National Academy of Sciences* **108,** 7265–7270 (2011).

7. Kaczynski, T., Mischaikow, K. & Mrozek, M. *Computational Homology* 1st ed. ISBN: 978-0-387-40853-8. doi:`10.1007/b97315` (Springer-Verlag New York, 2004).

8. Mischaikow, K. in *Dynamical Systems* (ed Johnson, R.) 119–207 (Springer Berlin Heidelberg, 1995). ISBN: 978-3-540-60047-3. doi:`10.1007/BFb0095240`.

9. De Silva, V. & Ghrist, R. Homological Sensor Networks. *Notices of the American Mathematical Society* **54,** 10–17 (2007).

10. Peacock, T., Froyland, G. & Haller, G. Introduction to Focus Issue: Objective Detection of Coherent Structures. *Chaos* **25,** 087201 (2015).

11. Dabiri, J. O. Optimal Vortex Formation as a Unifying Principle in Biological Propulsion. *Annual Review of Fluid Mechanics* **41,** 17–33 (2009).

12. Shadden, S. C. & Taylor, C. A. Characterization of Coherent Structures in the Cardiovascular System. *Annals of Biomedical Engineering* **36,** 1152–1162 (2008).

13. Haller, G. & Yuan, G.-C. Lagrangian coherent structures and mixing in two-dimensional turbulence. *Physica D: Nonlinear Phenomena* **147.** doi:`10.1016/S0167-2789(00)00142-1` (2000).

14. Haller, G. Lagrangian Coherent Structures. *Annual Review of Fluid Mechanics* **47.** doi:`10.1146/annurev-fluid-010313-141322` (2015).

15. Shadden, S. C., Lekien, F. & Marsden, J. E. Definition and properties of Lagrangian coherent structures from finite-time Lyapunov exponents in two-dimensional. *Physica D* **212,** 271–304 (2005).

16. Haller, G. A variational theory of hyperbolic Lagrangian Coherent Structures. *Physica D* **240,** 574–598 (2011).

17. Farazmand, M. & Haller, G. Erratum and addendum to 'A variational theory of hyperbolic Lagrangian coherent structures' [Physica D 240 (2011) 574-598]. *Physica D* **241,** 439–441 (2012).

18. Farazmand, M. & Haller, G. Computing Lagrangian coherent structures from their variational theory. *Chaos* **22,** 013128 (2012).

19. Munch, E. & Wang, B. Convergence between Categorical Representations of Reeb Space and Mapper (2016).

20. De Silva, V., Munch, E. & Patel, A. Categorified Reeb Graphs. *Discrete & Computational Geometry,* 1–53 (2016).

21. Carlsson, G. & Mémoli, F. Characterization, Stability and Convergence of Hierarchical Clustering Methods. *Journal of Machine Learning Research* **11,** 1425–1470 (2010).

22. *Arnold: Swimming Against the Tide* (eds Khesin, B. A. & Tabachnikov, S. L.) (American Mathematical Society, 2014).

23. Arnold, V. I. *Mathematical Methods of Classical Mechanics* trans. by Vogtmann, K. & Weinstein, A. doi:`10.1007/978-1-4757-2063-1` (Springer, 1989).

24. Lukatskii, A. Curvature of groups of diffeomorphisms preserving the measure of the 2–sphere. *Functional Analysis and Its Applications* **13,** 174–177 (1979).

25. Arnold, V. I. Sur la géométrie differérentielle des groupes de Lie de dimension infinie et ses applications a l'hydrodynamique des fluides parfaits. *Ann. Inst. Fourier* **16,** 319–361 (1966).

26. Aref, H. & Pomphrey, N. Integrable and chaotic motions of four vortices. *Physics Letters A* **78,** 297–300 (1980).

27. Aref, H. & Pomphrey, N. Integrable and Chaotic Motions of Four Vortices I. The Case of Identical Vortices. *Proc. Roy. Soc. Lond. A* **380,** 359–387 (1982).

28. Aref, H. Stirring by chaotic advection. *J. Fluid Mech.* **143,** 1–21 (1984).

29. Boyland, P., Stremler, M. & Aref, H. Topological fluid mechanics of point vortex motions. *Physica D: Nonlinear Phenomena* **175,** 69–95 (2003).

30. Boyland, P., Aref, H. & Stremler, M. Topological fluid mechanics of stirring. *J. Fluid Mech.* **403,** 277–304 (2000).

31. Gouillart, E., Thiffeault, J.-L. & Finn, M. D. Topological Mixing with ghost rods. *Phys. Rev. E* **73,** 036311 (2006).

32. Bowen, R. in *The Structure of Attractors in Dynamical Systems: Proceedings, North Dakota State University, June 20–24, 1977* (eds Markley, N. G., Martin, J. C. & Perrizo, W.) 21–29 (Springer, 1978). doi:`10.1007/BFb0101777`.

33. Thiffeault, J.-L. Measuring Topological Chaos. *Phys. Rev. Lett.* **94,** 084502 (2005).

34. Fried, D. Entropy and twisted cohomology. *Topology* **25,** 455–4679 (1985).

35. Kolev, B. Entropie topologique et représentation de Burau. *C.R. Acad. Sci. Paris Sér. I Math.* **309,** 835–838 (1989).

36. Allshouse, M. R. & Thiffeault, J.-L. Detecting coherent structures using braids. *Physica D: Nonlinear Phenomena* **241,** 95–105 (2012).

37. Bernardete, D., Gutierrez, M. & Nitecki, Z. in *Mapping Class Groups and Moduli Spaces of Riemann Surfaces* (eds Bodigheimer, C.-F. & Hain, R. M.) 1–31 (Amer. Math. Soc., 1993). doi:`10.1090/conm/150/01283`.

38. Bernardete, D., Nitecki, Z. & Gutierrez, M. Braids and the Nielsen-Thurston Classification. *J. Knot Theory Ramifications* **4.** doi:`10.1142/S0218216595000259` (1995).

39. Birman, J. S. & Brendle, T. E. in *Handbook of Knot Theory* (eds Menasco, W. & Thistlethwaite, M.) (Elsevier, 2005).

40. Farb, B. & Margalit, D. *A Primer on Mapping Class Groups* (Princeton University Press, 2012).

41. Boyland, P. Topological methods in surface dynamics. *Topology Appl.* **58,** 223–298 (1994).

42. Thiffeault, J.-L. Braids of entangled particle trajectories. *Chaos* **20,** 017516 (2010).

43. Budišić, M. & Thiffeault, J.-L. braidlab: a software package for braids and loops (2015).

44. Artin, E. Theory of braids. *Ann. of Math.* **48,** 101–126 (1947).

45. Thurston, W. P. & Levy, S. *Three-dimensional Geometry and Topology* (Princeton University Press, 1997).

46. Fathi, A., Laudenbach, F. & Poénaru, V. *Thurston's Work on Surfaces* trans. by Kim, D. M. & Margalit, D. (Princeton University Press, 2012).

47. Lee, E.-K. & Lee, S.-J. A Garside-theoretic approach to the reducibility problem in braid groups. *Journal of Algebra* **320,** 783–820 (2008).

48. Birman, J. S., Lubotzky, A. & McCarthy, J. Abelian and solvable subgroups of the mapping class groups. *Duke Math J.* **50,** 1107–1120 (1983).

49. Birman, J. S. *Braids, Links, and Mapping Class Groups* (Princeton University Press, 1975).

50. Long, D. & Paton, M. The Burau representation is not faithful for $n \geq 6$. *Topology* **32,** 439–447 (1993).

51. Bigelow, S. The Burau representation is not faithful for $n = 5$. *Geometry & Topology* **3,** 397–404 (1999).

52. Church, T. & Farb, B. Infinite generation of the kernels of the Magnus and Burau representations. *Algebraic & Geometric Topology* **10,** 837–851 (2010).

53. Birman, J., Long, D. & Moody, J. in *The Mathematical Legacy of Wilhelm Magnus: Groups, Geometry, and Special Functions* (eds Abikoff, Birman & Kuiken) (Amer. Math. Soc., 1992).

54. Rolfsen, D. *Tutorial on the Braid Groups* (eds Berrick, A. J., Cohen, F. R., Hanbury, E., Wong, Y.-L. & Wu, J.) ISBN: 978–9814291408 (World Scientific, Singapore, 2010).

55. Moody, J. The Burau representation of the Braid Group B_n is unfaithful for large n. *Bulletin of the American Mathematical Society* **25,** 379–384 (1991).

56. Band, G. & Boyland, P. The Burau estimate for the entropy of a braid. *Algebraic & Geometric Topology* **7,** 1345–1378 (2007).

57. González-Meneses, J. The nth root of a braid is unique up to conjugacy. *Algebraic & Geometric Topology* **3,** 1103–1118 (2003).

58. Ottino, J. *The kinematics of mixing: stretching, chaos, and transport* (Cambridge University Press, 1989).

59. Doherty, M. F. & Ottino, J. M. Chaos in deterministic systems: strange attractors, turbulence, and applications in chemical engineering. *Chemical Engineering Science* **43,** 139–183 (1988).

60. Strogatz, S. H. *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering* ISBN: 9780738204536 (Westview Press, Cambridge, MA, 2000).

61. Ruelle, D. & Takens, F. On the nature of turbulence. *Communications in Mathematical Physics* **20,** 167–192 (1971).

62. Milnor, J. On the concept of attractor. *Communications in Mathematical Physics* **99,** 177–195 (1985).

63. Bigelow, S. in *Proceedings of Symposia in Pure Mathematics* (eds Matić, G. & McCrory, C.) (2003). ISBN: 978-0-8218-3507-4. doi:`10.1090/pspum/071`.

64. Hu, T. C. & Shing, M. T. Computation of Matrix Chain Products. Part I. *SIAM Journal on Computing* **11,** 362–373 (1982).

65. Hu, T. C. & Shing, M. T. Computation of Matrix Chain Products. Part II. *SIAM J. Comput.* **13,** 228–251. ISSN: 0097-5397 (May 1984).

66. Lum, P. *et al.* Extracting insights from the shape of complex data using topology. *Scientific Reports* **3,** 1236 (2013).

67. Taylor, D. *et al.* Topological data analysis of contagion maps for examining spreading processes on networks. *Nature Communications* **3,** 7723 (2015).

68. Adams, H. & Carlsson, G. Evasion paths in mobile sensor networks. *The International Journal of Robotics Research* **34,** 90–104 (2014).

69. Niyogi, P., Smale, S. & Weinberger, S. Finding the homology of submanifolds with high confidence from random samples. *Discrete Comput Geom* **39,** 419–441 (2008).

70. Edelsbrunner, H., Letscher, D. & Zomorodian, A. *Topological persistence and simplification* in *Proceedings of the 41st Annual Symposium on Foundations of Computer Science* (2000), 454–. <`http://dl.acm.org/citation.cfm?id=795666.796607`>.

71. Carlsson, G., de Silva, V. & Morozov, D. *Zigzag persistent homology and real-valued functions* in *Proceedings of the twenty-fifth annual symposium on Computational geometry* (ACM, Aarhus, Denmark, 2009), 247–256. ISBN: 978-1-60558-501-7. doi:`10.1145/1542362.1542408`.

72. Munkres, J. R. *Elements of Algebraic Topology* ISBN: 0-201-62728-0 (Perseus Publishing, Cambridge, Massachusetts, 1984).

73. Hatcher, A. Algebraic Topology (2002).

74. Borsuk, K. On the imbedding of systems of compacta in simplicial complexes. *Fund. Math* **35,** 217–234 (1948).

75.   Björner, A. Nerves, fibers and homotopy groups. *Journal of Combinatorial Theory, Series A* **102,** 88–93 (2003).

76.   Carlsson, G. Topology and data. *Bull. Amer. Math. Soc.* **46,** 255–308 (2009).

77.   Edelsbrunner, H. & Harer, J. in *Surveys on Discrete and Computational Geometry: Twenty Years Later* (eds Goodman, J. E., Pach, J. & Pollack, R.) 257–282 (American Mathematical Society, 2008). doi:`10.1090/conm/453/08802`.

78.   Edelsbrunner, H. & Harer, J. *Computational Topology: An Introduction* (American Mathematical Society, Providence, R.I., 2010).

79.   Reeb, G. Sur les points singuliers d'une forme de Pfaff complèment intégrable ou d'une fonction numérique. *Comptes Rendus de L'Académie des Sciences* **222,** 847–849 (1946).

80.   Hilaga, M., Shinagawa, Y., Kohmura, T. & Kunii, T. L. *Topology matching for fully automatic similarity estimation of 3D shapes* in *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques* (ACM, New York, NY, USA, 2001), 203–212. ISBN: 1-58113-374-X. doi:`10.1145/383259.383282`.

81.   Escolano, F., Hancock, E. R. & Biasotti, S. in *Computer Analysis of Images and Patterns* (eds Wilson, R., Hancock, E., Bors, A. & Smith, W.) 120–127 (Springer Berlin Heidelberg, 2013).

82.   Ge, X., Safa, I. I., Belkin, M. & Wang, Y. in *Advances in Neural Information Processing Systems 24* (eds Shawe-Taylor, J., Zemel, R. S., Bartlett, P. L., Pereira, F. & Weinberger, K. Q.) 837–845 (Curran Associates, Inc., 2011). <http://papers.nips.cc/paper/4375-data-skeletonization-via-reeb-graphs.pdf>.

83.   Wood, Z., Hoppe, H., Desbrun, M. & Schröder, P. Removing Excess Topology from Isosurfaces. *ACM Trans. Graph.* **23,** 190–208. ISSN: 0730-0301 (Apr. 2004).

84.   Pascucci, V., Scorzelli, G., Bremer, P.-T. & Mascarenhas, A. Robust Online Computation of Reeb Graphs: Simplicity and Speed. *ACM Trans. Graph.* **26.** ISSN: 0730-0301. doi:`10.1145/1276377.1276449` (2007).

85.   Pascucci, V., Scorzelli, G., Bremer, P.-T. & Mascarenhas, A. *Robust Online Computation of Reeb Graphs: Simplicity and Speed* in *ACM SIGGRAPH 2007 Papers* (ACM, San Diego, California, 2007). doi:`10.1145/1275808.1276449`.

86.   Carlsson, G. & Mémoli, F. Classifying Clustering Schemes. *Foundations of Computational Mathematics* **13,** 221–252 (2013).

87. Hamza, A. B. & Krim, H. in *Discrete Geometry for Computer Imagery* (eds Nyström, I., di Baja, G. S. & Svensson, S.) 378–387 (Springer Berlin Heidelberg, 2003). doi:`10.1007/978-3-540-39966-7_36`.

88. Chazal, F., Cohen-Steiner, D., Glisse, M. & Guibas, L. Proximity of Persistence Modules and Their Diagrams. *Symposium on Computational Geometry '09* 237–246 (2009).

89. Reynolds, R. W. *et al.* Daily High-Resolution-Blended Analyses for Sea Surface Temperature. *Journal of Climate* **20,** 5473–5496.

90. Morozov, D. Dionysus: a library for computing persistent homology (2007).